

FILE COPY

Receipt



UNITED STATES PATENT AND TRADEMARK OFFICE

COMMISSIONER FOR PATENTS
UNITED STATES PATENT AND TRADEMARK OFFICE
WASHINGTON, D.C. 20231
www.uspto.gov

APPLICATION NUMBER	FILING DATE	GRP ART UNIT	FIL FEE REC'D	ATTY. DOCKET NO	DRAWINGS	TOT CLAIMS	IND CLAIMS
09/619,411	07/19/2000	2761	838	108.0001	25	21	3

FILING RECEIPT



OC000000005504037

Peter H Priest
Law Offices of Peter H Priest
529 Dogwood Drive
Chapel Hill, NC 27516

Date Mailed: 10/25/2000

Receipt is acknowledged of this nonprovisional Patent Application. It will be considered in its order and you will be notified as to the results of the examination. Be sure to provide the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION when inquiring about this application. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. If an error is noted on this Filing Receipt, please write to the Office of Initial Patent Examination's Customer Service Center. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the PTO processes the reply to the Notice, the PTO will generate another Filing Receipt incorporating the requested corrections (if appropriate).

Applicant(s)

Jeffrey Lynn Eakin, Ridgefield, CT ;
Christopher Huyette Howe, Stamford, CT ;
Joseph Eric Lipovich, Trumbull, CT ;
Mark Wesley Brewster, NY ;
Michael James Platt, Fairfield, CT ;

Mark Wesley MacKenzie

Technology Center 2100

FEB 12 2001

RECEIVED

Continuing Data as Claimed by Applicant

Foreign Applications

If Required, Foreign Filing License Granted 09/12/2000

Title

Methods and apparatus for processing and distributing information relating to costs and sales of products

Preliminary Class

705

Third Try
Please Correct

Data entry by : ARMSTEAD, LORRAINE

Team : OIPE

Date: 10/25/2000



**LICENSE FOR FOREIGN FILING UNDER
Title 35, United States Code, Section 184
Title 37, Code of Federal Regulations, 5.11 & 5.15**

**GRANTED**

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 36 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Office of Export Administration, Department of Commerce (15 CFR 370.10 (j)); the Office of Foreign Assets Control, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

PLEASE NOTE the following information about the Filing Receipt:

- The articles such as "a," "an" and "the" are not included as the first words in the title of an application. They are considered to be unnecessary to the understanding of the title.
- The words "new," "improved," "improvements in" or "relating to" are not included as first words in the title of an application because a patent application, by nature, is a new idea or improvement.
- The title may be truncated if it consists of more than 600 characters (letters and spaces combined).
- The docket number allows a maximum of 25 characters.
- If your application was submitted under 37 CFR 1.10, your filing date should be the "date in" found on the Express Mail label. If there is a discrepancy, you should submit a request for a corrected Filing Receipt along with a copy of the Express Mail label showing the "date in."
- The title is recorded in sentence case.

Any corrections that may need to be done to your Filing Receipt should be directed to:

Assistant Commissioner for Patents
Office of Initial Patent Examination
Customer Service Center
Washington, DC 20231

RECEIVED
FEB 12 2001
Technology Center 2100



FILE COPY

UNITED STATES PATENT AND TRADEMARK OFFICE

 COMMISSIONER FOR PATENTS
 UNITED STATES PATENT AND TRADEMARK OFFICE
 WASHINGTON, D.C. 20231
 www.uspto.gov


Bib Data Sheet

SERIAL NUMBER 09/619,411	FILING DATE 07/19/2000 RULE -	CLASS 705	GROUP ART UNIT 2761 2161	ATTORNEY DOCKET NO. 108.0001
APPLICANTS Jeffrey Lynn Eakin, Ridgefield, CT ; Christopher Huyette Howe, Stamford, CT ; Joseph Eric Lipovich, Trumbull, CT ; Mark Wesley Mackenzie, Brewster, NY ; Michael James Platt, Fairfield, CT ;				
** CONTINUING DATA ***** ** FOREIGN APPLICATIONS *****				
IF REQUIRED, FOREIGN FILING LICENSE GRANTED ** 09/12/2000				
Foreign Priority claimed <input type="checkbox"/> yes <input type="checkbox"/> no 35 USC 119 (a-d) conditions <input type="checkbox"/> yes <input type="checkbox"/> no <input type="checkbox"/> Met after Allowance Verified and Acknowledged _____ Examiner's Signature Initials		STATE OR COUNTRY CT	SHEETS DRAWING 25	TOTAL CLAIMS 21
				INDEPENDENT CLAIMS 3
ADDRESS Peter H Priest Law Offices of Peter H Priest 529 Dogwood Drive Chapel Hill, NC 27516				
TITLE Methods and apparatus for processing and distributing information relating to costs and sales of products				
FILING FEE RECEIVED 838	FEES: Authority has been given in Paper No. _____ to charge/credit DEPOSIT ACCOUNT No. _____ for following:		<input type="checkbox"/> All Fees <input type="checkbox"/> 1.16 Fees (Filing) <input type="checkbox"/> 1.17 Fees (Processing Ext. of time) <input type="checkbox"/> 1.18 Fees (Issue) <input type="checkbox"/> Other _____ <input type="checkbox"/> Credit	

 RECEIVED
 FEB 12 2001
 Technology Center 210

CC without a sequence coding for SP; (2) an isolated DNA molecule (IV)
 CC comprising at least a part of (III) and (I) that is functional in a
 CC lactic acid bacterium; (3) an isolated DNA sequence (I) coding for SP
 CC derived from SP10, SP13, SP307, SP310 or SP330, or a derivative of any of
 CC the original peptides having retained signal peptide functionality; (4) a
 CC recombinant plasmid (V) comprising (II) or (I); and (5) a recombinant
 CC bacterium (VI) comprising (I). (III) is useful for identifying and
 CC isolating (I) from a source lactic acid bacteria, by transforming the
 CC bacteria with (III), and selecting from the transformed bacteria, cells
 CC in which the promoterless promoter reporter gene is expressed and the
 CC gene product of the DNA sequence coding for a secretion reporter molecule
 CC is secreted. (VI) is useful for the production of a desired gene product.
 CC AAF59460 to AAF59499 and AAB70428 to AAB70472 represent sequences used in
 CC the exemplification of the present invention

XX Sequence 36 AA;

Query Match 90.0%; Score 148.5; DB 4; Length 36;
 Best Local Similarity 94.4%; Pred. No. 4,4e-15;
 Matches 34; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 1 MKENKRVATATFIALIFVSFTTIS-ODQAQAERS 35
 |||||
 1 MKENKRVATATFIALIFVSFTTISODQAQAERS 36

RESULT 8

AAB70437
 ID AAB70437 standard; peptide; 36 AA.

AC AAB70437;

DT 03-MAY-2001 (first entry)

DE L. lactis signal peptide SP310 mutant amino acid sequence SEQ:50.

XX Lactococcus lactis; signal peptide; secretion signal; SP310; mutant;
 KM lactic acid bacterium; promoterless promoter reporter; PPR; transposon;
 XX identification; secretion reporter.

OS Lactococcus lactis.
 OS Synthetic.

PN WO200111060-A2.

PD 15-FEB-2001.

PP 04-AUG-2000; 2000WO-DK000437.

PR 06-AUG-1999; 99DK-00001105.

PA (BIOT-) BIOTEKNOLOGISK INST.

PI Ravn P, Madsen SM, Vrang A, Israeljen H, Johnsen MG, Bredmose L,
 PI Arnanu J,

XX WPI; 2001-191547/19.

PT Constructing a transposon derivative to identify DNA sequence encoding
 PT signal peptide in lactic acid bacteria, involves removing stop codons in
 PT frame with secretion reporter molecule from DNA comprising transposon.

PS Example 2; Page 60; 62pp; English.

XX The present invention describes a method for constructing a transposon
 CC derivative for identifying DNA (I) encoding a signal peptide (secretion
 CC signal, SP) in a lactic acid bacterium (e.g. Lactococcus lactis). The
 CC method comprises selecting a transposon (II), including a promoterless
 CC promoter reporter (PPR) gene and a ribosome binding site (RBS), between
 CC its left and right termini (LR) and (RR), deleting a region between LR
 CC and PPR gene to obtain modified DNA that retains its transposability and
 CC its RBS. The present invention also describes: (1) a transposon
 CC derivative (III) for the identification of (I) in a lactic acid

CC bacterium, comprising (II) without stop codons in the region upstream of
 CC the PPR gene, and a DNA sequence encoding a secretion reporter molecule
 CC without a sequence coding for SP; (2) an isolated DNA molecule (IV)
 CC comprising at least a part of (III) and (I) that is functional in a
 CC lactic acid bacterium; (3) an isolated DNA sequence (I) coding for SP
 CC derived from SP10, SP13, SP307, SP310 or SP330, or a derivative of any of
 CC the signal peptides having retained signal peptide functionality; (4) a
 CC recombinant plasmid (V) comprising (II) or (I); and (5) a recombinant
 CC bacterium (VI) comprising (I). (III) is useful for identifying and
 CC isolating (I) from a source lactic acid bacteria, by transforming the
 CC bacteria with (III), and selecting from the transformed bacteria, cells
 CC in which the promoterless promoter reporter gene is expressed and the
 CC gene product of the DNA sequence coding for a secretion reporter molecule
 CC is secreted. (VI) is useful for the production of a desired gene product.
 CC AAF59460 to AAF59499 and AAB70428 to AAB70472 represent sequences used in
 CC the exemplification of the present invention

SQ Sequence 36 AA;

Query Match 89.4%; Score 147.5; DB 4; Length 36;
 Best Local Similarity 94.4%; Pred. No. 6,2e-15;
 Matches 34; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 1 MKENKRVATATFIALIFVSFTTIS-ODQAQAERS 35
 |||||
 1 MKENKRVATATFIALIFVSFTTISODQAQAERS 36

RESULT 9

AAB70434
 ID AAB70434 standard; peptide; 36 AA.

AC AAB70434;

DT 03-MAY-2001 (first entry)

DE L. lactis signal peptide SP310 mutant amino acid sequence SEQ:47.

XX Lactococcus lactis; signal peptide; secretion signal; SP310; mutant;
 KM lactic acid bacterium; promoterless promoter reporter; PPR; transposon;
 XX identification; secretion reporter.

OS Lactococcus lactis.
 OS Synthetic.

PN WO200111060-A2.

PD 15-FEB-2001.

PP 04-AUG-2000; 2000WO-DK000437.

PR 06-AUG-1999; 99DK-00001105.

PA (BIOT-) BIOTEKNOLOGISK INST.

PI Ravn P, Madsen SM, Vrang A, Israeljen H, Johnsen MG, Bredmose L,
 PI Arnanu J,

XX WPI; 2001-191547/19.

PT Constructing a transposon derivative to identify DNA sequence encoding
 PT signal peptide in lactic acid bacteria, involves removing stop codons in
 PT frame with secretion reporter molecule from DNA comprising transposon.

PS Example 2; Page 59; 62pp; English.

XX The present invention describes a method for constructing a transposon
 CC derivative for identifying DNA (I) encoding a signal peptide (secretion
 CC signal, SP) in a lactic acid bacterium (e.g. Lactococcus lactis). The
 CC method comprises selecting a transposon (II), including a promoterless
 CC promoter reporter (PPR) gene and a ribosome binding site (RBS), between
 CC its left and right termini (LR) and (RR), deleting a region between LR
 CC and PPR gene to obtain modified DNA that retains its transposability and

CC its RBS. The present invention also describes: (1) a transposon
 CC derivative (II) for the identification of (I) in a lactic acid
 CC bacterium, comprising (II) without stop codons in the region upstream of
 CC the PPR gene, and a DNA sequence encoding a secretion reporter molecule
 CC without a sequence coding for SP; (2) an isolated DNA molecule (IV)
 CC comprising at least a part of (III) and (I) that is functional in a
 CC lactic acid bacterium; (3) an isolated DNA sequence (I) coding for SP
 CC derived from SP10, SP13, SP307, SP310 or SP330, or a derivative of any of
 CC the signal peptides having retained signal peptide functionality; (4) a
 CC recombinant plasmid (V) comprising (II) or (I); and (5) a recombinant
 CC bacterium (VI) comprising (I), (II) is useful for identifying and
 CC isolating (I) from a source lactic acid bacteria, by transforming the
 CC bacteria with (III), and selecting from the transformed bacteria, cells
 CC in which the promoterless promoter reporter gene is expressed and the
 CC gene product of the DNA sequence coding for a secretion reporter molecule
 CC is secreted. (VI) is useful for the production of a desired gene product.
 CC AAF59460 to AAF59499 and AAB70428 to AAB70472 represent sequences used in
 CC the exemplification of the present invention

CC Sequence 36 AA;

Query Match Best Local Similarity 91.7%; Score 142.5; DB 4; Length 36;
 Matches 33; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 MKEFKKRAIATPFIATFIIVSFPTISS-ODQAQAERS 35
 DB 1 MKEFKKRAIATPFIATFIIVSFPTISSIQDAQAEDTRS 36

RESULT 10

-ID AAB70439 standard; peptide; 36 AA.

AC AAB70439;

DT 03-MAY-2001 (first entry)

DE L. lactis signal peptide SP310 mutant amino acid sequence SEQ:52.

KW Lactococcus lactis; signal peptide; secretion signal; SP310; mutant;
 KW lactic acid bacterium; promoterless promoter reporter; PPR; transposon;
 OS identification; secretion reporter.

OS Lactococcus lactis.
 OS Synthetic.

PN WO200111060-A2.

PD 15-FEB-2001.

PF 04-AUG-2000; 2000WO-DK000437.

PR 06-AUG-1999; 99DK-00001105.

PA (BIOT-) BIOTEKNOLOGISK INST.

PI Ravn P, Madsen SM, Vrang A, Israelsen H, Johnsen MG, Bredmose L;
 PI Arnau J;

PI WPI; 2001-191547/19.

PT Constructing a transposon derivative to identify DNA sequence encoding
 PT signal peptide in lactic acid bacteria, involves removing stop codons in
 PT frame with secretion reporter molecule from DNA comprising transposon.

PS Example 2; Page 60; 62pp; English.

CC The present invention describes a method for constructing a transposon
 CC derivative for identifying DNA (I) encoding a signal peptide (secretion
 CC signal, SP) in a lactic acid bacterium (e.g. Lactococcus lactis). The
 CC method comprises selecting a transposon (II), including a promoterless
 CC promoter reporter (PPR) gene and a ribosome binding site (RBS), between

CC its left and right termini (LR) and (RR), deleting a region between LR
 CC and PPR gene to obtain modified DNA that retains its transposability and
 CC its RBS. The present invention also describes: (1) a transposon
 CC derivative (II) for the identification of (I) in a lactic acid
 CC bacterium, comprising (II) without stop codons in the region upstream of
 CC the PPR gene, and a DNA sequence encoding a secretion reporter molecule
 CC without a sequence coding for SP; (2) an isolated DNA molecule (IV)
 CC comprising at least a part of (III) and (I) that is functional in a
 CC lactic acid bacterium; (3) an isolated DNA sequence (I) coding for SP
 CC derived from SP10, SP13, SP307, SP310 or SP330, or a derivative of any of
 CC the signal peptides having retained signal peptide functionality; (4) a
 CC recombinant plasmid (V) comprising (II) or (I); and (5) a recombinant
 CC bacterium (VI) comprising (I), (II) is useful for identifying and
 CC isolating (I) from a source lactic acid bacteria, by transforming the
 CC bacteria with (III), and selecting from the transformed bacteria, cells
 CC in which the promoterless promoter reporter gene is expressed and the
 CC gene product of the DNA sequence coding for a secretion reporter molecule
 CC is secreted. (VI) is useful for the production of a desired gene product.
 CC AAF59460 to AAF59499 and AAB70428 to AAB70472 represent sequences used in
 CC the exemplification of the present invention

CC Sequence 36 AA;

Query Match Best Local Similarity 91.7%; Score 142.5; DB 4; Length 36;
 Matches 33; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 MKEFKKRAIATPFIATFIIVSFPTISS-ODQAQAERS 35
 DB 1 MKEFKKRAIATPFIATFIIVSFPTISSIQDAQAERS 36

RESULT 11

ID AAB70431 standard; peptide; 35 AA.

AC AAB70431;

DT 03-MAY-2001 (first entry)

DE L. lactis signal peptide SP310 mutant amino acid sequence SEQ:44.

KW Lactococcus lactis; signal peptide; secretion signal; SP310; mutant;
 KW lactic acid bacterium; promoterless promoter reporter; PPR; transposon;
 OS identification; secretion reporter.

OS Lactococcus lactis.
 OS Synthetic.

PN WO200111060-A2.

PD 15-FEB-2001.

PF 04-AUG-2000; 2000WO-DK000437.

PR 06-AUG-1999; 99DK-00001105.

PA (BIOT-) BIOTEKNOLOGISK INST.

PI Ravn P, Madsen SM, Vrang A, Israelsen H, Johnsen MG, Bredmose L;
 PI Arnau J;

PI WPI; 2001-191547/19.

PT Constructing a transposon derivative to identify DNA sequence encoding
 PT signal peptide in lactic acid bacteria, involves removing stop codons in
 PT frame with secretion reporter molecule from DNA comprising transposon.

PS Example 2; Page 58; 62pp; English.

CC The present invention describes a method for constructing a transposon
 CC derivative for identifying DNA (I) encoding a signal peptide (secretion
 CC signal, SP) in a lactic acid bacterium (e.g. Lactococcus lactis). The

CC method comprises selecting a transposon (II), including a promoterless
 CC promoter reporter (PPR) gene and a ribosome binding site (RBS), between
 CC its left and right termini (LR) and (RR), deleting a region between LR
 CC and PPR gene to obtain modified DNA that retains its transposability and
 CC its RBS. The present invention also describes: (1) a transposon
 CC derivative (III) for the identification of (I) in a lactic acid
 CC bacterium, comprising (II) without stop codons in the region upstream of
 CC the PPR gene, and a DNA sequence encoding a secretion reporter molecule
 CC without a sequence coding for SP; (2) an isolated DNA molecule (IV)
 CC comprising at least a part of (III) and (I) that is functional in a
 CC lactic acid bacterium; (3) an isolated DNA sequence (I) coding for SP
 CC derived from SP10, SP13, SP307, SP310 or SP330, or a derivative of any of
 CC the signal peptides having retained signal peptide functionality; (4) a
 CC recombinant plasmid (V) comprising (II) or (I); and (5) a recombinant
 CC bacterium (VI) comprising (I). (III) is useful for identifying and
 CC isolating (I) from a source lactic acid bacteria, by transforming the
 CC bacteria with (III), and selecting from the transformed bacteria, cells
 CC in which the promoterless promoter reporter gene is expressed and the
 CC gene product of the DNA sequence coding for a secretion reporter molecule
 CC is secreted. (VI) is useful for the production of a desired gene product.
 CC AAF59460 to AAF59499 and AAB70428 to AAB70472 represent sequences used in
 CC the exemplification of the present invention

XX Sequence 35 AA;

Query Match Best Local Similarity 86.1%; Score 142; DB 4; Length 35;

Matches 31; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKFNKKRAVAATFATFALFVSPFTTSSQDAQAARS 35
 |||||
 1 MKFNKKRAVAATFATFALFVSPFTTINTQAARS 35

RESULT 12

AAB70446 standard; peptide; 34 AA.

XX AAB70446;

DT 03-MAY-2001 (first entry)

DE L. lactis signal peptide SP310 mutant amino acid sequence SEQ:59.

XX Lactococcus lactis; signal peptide; secretion signal; SP310; mutant;

KM lactic acid bacterium; promoterless promoter reporter; PPR; transposon;

XX identification; secretion reporter.

OS Lactococcus lactis.

XX Synthetic.

PN WO200111060-A2.

XX 15-FEB-2001.

PE 04-AUG-2000; 2000WO-DK000437.

PR 06-AUG-1999; 99DK-00001105.

XX (BIOT-) BIOTEKNOLOGISK INST.

PI Ravn P, Madsen SM, Vrang A, Israelsen H, Johnsen MG, Bredmose L,
 PI Arnau J;

XX WPI; 2001-191547/19.

PT Constructing a transposon derivative to identify DNA sequence encoding
 PT signal peptide in lactic acid bacteria, involves removing stop codons in
 PT frame with secretion reporter molecule from DNA comprising transposon.
 XX Example 2; Page 62; 62pp; English.

XX The present invention describes a method for constructing a transposon

CC derivative for identifying DNA (I) encoding a signal peptide (secretion
 CC signal, SP) in a lactic acid bacterium (e.g. Lactococcus lactis). The
 CC method comprises selecting a transposon (II), including a promoterless
 CC promoter reporter (PPR) gene and a ribosome binding site (RBS), between
 CC its left and right termini (LR) and (RR), deleting a region between LR
 CC and PPR gene to obtain modified DNA that retains its transposability and
 CC its RBS. The present invention also describes: (1) a transposon
 CC derivative (III) for the identification of (I) in a lactic acid
 CC bacterium, comprising (II) without stop codons in the region upstream of
 CC the PPR gene, and a DNA sequence encoding a secretion reporter molecule
 CC without a sequence coding for SP; (2) an isolated DNA molecule (IV)
 CC comprising at least a part of (III) and (I) that is functional in a
 CC lactic acid bacterium; (3) an isolated DNA sequence (I) coding for SP
 CC derived from SP10, SP13, SP307, SP310 or SP330, or a derivative of any of
 CC the signal peptides having retained signal peptide functionality; (4) a
 CC recombinant plasmid (V) comprising (II) or (I); and (5) a recombinant
 CC bacterium (VI) comprising (I). (III) is useful for identifying and
 CC isolating (I) from a source lactic acid bacteria, by transforming the
 CC bacteria with (III), and selecting from the transformed bacteria, cells
 CC in which the promoterless promoter reporter gene is expressed and the
 CC gene product of the DNA sequence coding for a secretion reporter molecule
 CC is secreted. (VI) is useful for the production of a desired gene product.
 CC AAF59460 to AAF59499 and AAB70428 to AAB70472 represent sequences used in
 CC the exemplification of the present invention

XX Sequence 34 AA;

Query Match Best Local Similarity 85.5%; Score 141; DB 4; Length 34;

Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 NKKRAVAATFATFALFVSPFTTSSQDAQAARS 35
 |||||
 3 NKKRAVAATFATFALFVSPFTTSSQDAQAARS 34

DB 3

RESULT 13 AAB70433 standard; peptide; 35 AA.

XX AAB70433;

DT 03-MAY-2001 (first entry)

DE L. lactis signal peptide SP310 mutant amino acid sequence SEQ:46.

XX Lactococcus lactis; signal peptide; secretion signal; SP310; mutant;

KM lactic acid bacterium; promoterless promoter reporter; PPR; transposon;

XX identification; secretion reporter.

OS Lactococcus lactis.

PN WO200111060-A2.

XX 15-FEB-2001.

PE 04-AUG-2000; 2000WO-DK000437.

PR 06-AUG-1999; 99DK-00001105.

XX (BIOT-) BIOTEKNOLOGISK INST.

PI Ravn P, Madsen SM, Vrang A, Israelsen H, Johnsen MG, Bredmose L,
 PI Arnau J;

XX WPI; 2001-191547/19.

PT Constructing a transposon derivative to identify DNA sequence encoding
 PT signal peptide in lactic acid bacteria, involves removing stop codons in
 PT frame with secretion reporter molecule from DNA comprising transposon.
 XX Example 2; Page 59; 62pp; English.

XX The present invention describes a method for constructing a transposon

XX The present invention describes a method for constructing a transposon
 CC derivative for identifying DNA (I) encoding a signal peptide (secretion
 CC signal, SP) in a lactic acid bacterium (e.g. *Lactococcus lactis*). The
 CC method comprises selecting a transposon (II), including a promoterless
 CC promoter reporter (PPR) gene and a ribosome binding site (RBS), between
 CC its left and right termini (LR) and (RR), deleting a region between LR
 CC and PPR gene to obtain modified DNA that retains its transposability and
 CC its RBS. The present invention also describes: (1) a transposon
 CC derivative (III) for the identification of (I) in a lactic acid
 CC bacterium, comprising (II) without stop codons in the region upstream of
 CC the PPR gene, and a DNA sequence encoding a secretion reporter molecule
 CC without a sequence coding for SP; (2) an isolated DNA molecule (IV)
 CC comprising at least a part of (III) and (I) that is functional in a
 CC lactic acid bacterium; (3) an isolated DNA sequence (I) coding for SP
 CC derived from SP10, SP13, SP307, SP310 or SP330, or a derivative of any of
 CC the signal peptides having retained signal peptide functionality; (4) a
 CC recombinant plasmid (V) comprising (II) or (I); and (5) a recombinant
 CC bacterium (VI) comprising (I). (III) is useful for identifying and
 CC isolating (I) from a source lactic acid bacteria, by transforming the
 CC bacteria with (III), and selecting from the transformed bacteria, cells
 CC in which the promoterless promoter reporter gene is expressed and the
 CC gene product of the DNA sequence coding for a secretion reporter molecule
 CC is secreted. (VI) is useful for the production of a desired gene product.
 CC AAF59460 to AAF59499 and AAB70428 to AAB70472 represent sequences used in
 CC the exemplification of the present invention

XX Sequence 35 AA;
 SQ

Query Match 84.2%; Score 139; DB 4; Length 35;
 Best Local Similarity 85.7%; Pred. No. 1.2e-13;
 Matches 30; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKFNKKVAIATPFIATFIVSFPTISSQDAQAERS 35
 Db 1 MKFNKKVAIATPFIATFIVSFPTISSQDAQAERS 35

RESULT 14
 AAB70445
 ID AAB70445 standard; peptide; 35 AA.
 AC AAB70445;
 XX
 XX AAB70445;
 DT 03-MAY-2001 (first entry)
 XX
 DE L. lactis signal peptide SP310 mutant amino acid sequence SEQ:58.
 XX
 KW *Lactococcus lactis*; signal peptide; secretion signal; SP310; mutant;
 KW lactic acid bacterium; promoterless promoter reporter; PPR; transposon;
 KW identification; secretion reporter.
 XX
 OS *Lactococcus lactis*.
 OS Synthetic.
 OS
 OS WO200111060-A2.
 XX
 XX WO200111060-A2.
 PD 15-FEB-2001.
 XX
 PF 04-AUG-2000; 2000WO-DK000437.
 XX
 PR 06-AUG-1999; 99DK-00001105.
 XX
 PA (BIOT-) BIOTEKNOLOGISK INST.
 XX
 PI Ravn P, Madsen SM, Vrang A, Israelsen H, Johnsen MG, Bredmose L;
 PI Arnau J;
 XX
 XX WPI; 2001-191547/19.

Constructing a transposon derivative to identify DNA sequence encoding
 PT signal peptide in lactic acid bacteria, involves removing stop codons in
 PT frame with secretion reporter molecule from DNA comprising transposon.

XX Example 2; Page 62; 62pp; English.
 PS
 XX
 XX The present invention describes a method for constructing a transposon
 CC derivative for identifying DNA (I) encoding a signal peptide (secretion
 CC signal, SP) in a lactic acid bacterium (e.g. *Lactococcus lactis*). The
 CC method comprises selecting a transposon (II), including a promoterless
 CC promoter reporter (PPR) gene and a ribosome binding site (RBS), between
 CC its left and right termini (LR) and (RR), deleting a region between LR
 CC and PPR gene to obtain modified DNA that retains its transposability and
 CC its RBS. The present invention also describes: (1) a transposon
 CC derivative (III) for the identification of (I) in a lactic acid
 CC bacterium, comprising (II) without stop codons in the region upstream of
 CC the PPR gene, and a DNA sequence encoding a secretion reporter molecule
 CC without a sequence coding for SP; (2) an isolated DNA molecule (IV)
 CC comprising at least a part of (III) and (I) that is functional in a
 CC lactic acid bacterium; (3) an isolated DNA sequence (I) coding for SP
 CC derived from SP10, SP13, SP307, SP310 or SP330, or a derivative of any of
 CC the signal peptides having retained signal peptide functionality; (4) a
 CC recombinant plasmid (V) comprising (II) or (I); and (5) a recombinant
 CC bacterium (VI) comprising (I). (III) is useful for identifying and
 CC isolating (I) from a source lactic acid bacteria, by transforming the
 CC bacteria with (III), and selecting from the transformed bacteria, cells
 CC in which the promoterless promoter reporter gene is expressed and the
 CC gene product of the DNA sequence coding for a secretion reporter molecule
 CC is secreted. (VI) is useful for the production of a desired gene product.
 CC AAF59460 to AAF59499 and AAB70428 to AAB70472 represent sequences used in
 CC the exemplification of the present invention

XX Sequence 35 AA;
 SQ

Query Match 83.9%; Score 138.5; DB 4; Length 35;
 Best Local Similarity 97.0%; Pred. No. 1.4e-13;
 Matches 32; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 4 NKKRVAIATPFIATFIVSFPTISSQDAQAERS 35
 Db 3 NKKRVAIATPFIATFIVSFPTISSQDAQAERS 35

RESULT 15
 AAB70429
 ID AAB70429 standard; peptide; 38 AA.
 AC AAB70429;
 XX
 XX AAB70429;
 DT 03-MAY-2001 (first entry)
 XX
 DE L. lactis signal peptide SP310 mutant amino acid sequence SEQ:42.
 XX
 KW *Lactococcus lactis*; signal peptide; secretion signal; SP310; mutant;
 KW lactic acid bacterium; promoterless promoter reporter; PPR; transposon;
 KW identification; secretion reporter.
 XX
 OS *Lactococcus lactis*.
 OS Synthetic.
 OS
 OS WO200111060-A2.
 XX
 XX WO200111060-A2.
 PD 15-FEB-2001.
 XX
 PF 04-AUG-2000; 2000WO-DK000437.
 XX
 PR 06-AUG-1999; 99DK-00001105.
 XX
 PA (BIOT-) BIOTEKNOLOGISK INST.
 XX
 PI Ravn P, Madsen SM, Vrang A, Israelsen H, Johnsen MG, Bredmose L;
 PI Arnau J;
 XX
 XX WPI; 2001-191547/19.

Constructing a transposon derivative to identify DNA sequence encoding

PT signal peptide in lactic acid bacteria, involves removing stop codons in
 PT frame with secretion reporter molecule from DNA comprising transposon.

XX
 XX
 XX Example 2; Page 58; 62pp; English.

CC The present invention describes a method for constructing a transposon
 CC derivative for identifying DNA (I) encoding a signal peptide (secretion
 CC signal, SP) in a lactic acid bacterium (e.g. *Lactococcus lactis*). The
 CC method comprises selecting a transposon (II), including a promoterless
 CC promoter reporter (PPR) gene and a ribosome binding site (RBS), between
 CC its left and right termini (LR) and (RR), deleting a region between LR
 CC and PPR gene to obtain modified DNA that retains its transposability and
 CC its RBS. The present invention also describes: (1) a transposon
 CC derivative (III) for the identification of (I) in a lactic acid
 CC bacterium, comprising (II) without stop codons in the region upstream of
 CC the PPR gene, and a DNA sequence encoding a secretion reporter molecule
 CC without a sequence coding for SP; (2) an isolated DNA molecule (IV)
 CC comprising at least a part of (III) and (I) that is functional in a
 CC lactic acid bacterium; (3) an isolated DNA sequence (I) coding for SP
 CC derived from SP10, SP13, SP107, SP110 or SP130, or a derivative of any of
 CC the signal peptides having retained signal peptide functionality; (4) a
 CC recombinant plasmid (V) comprising (II) or (I); and (5) a recombinant
 CC bacterium (VI) comprising (I). (III) is useful for identifying and
 CC isolating (I) from a source lactic acid bacteria, by transforming the
 CC bacteria with (III), and selecting from the transformed bacteria, cells
 CC in which the promoterless promoter reporter gene is expressed and the
 CC gene product of the DNA sequence coding for a secretion reporter molecule
 CC is secreted. (VI) is useful for the production of a desired gene product.
 CC AAF59460 to AAF59499 and AAB70428 to AAB70472 represent sequences used in
 CC the exemplification of the present invention

XX
 XX SQ, sequence 38 AA;

Query Match 83.9%; Score 138.5; DB 4; Length 38;
 Best Local Similarity 89.5%; Pred. No. 1.5e-13;
 Matches 34; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY 1 MKFNKRVAIAIPFIALIFVSFTTSS-QD--AQAAERS 35
 |||
 DB 1 MKFNKRVAIAIPFIALIFVSFTTSSIQDNQANAERS 38

Search completed: November 7, 2004, 17:54:44
 Job time : 80 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2004, 14:58:19 ; Search time 25 Seconds
(without alignments)
92.845 Million cell updates/sec

Title: US-09-982-531-1

Perfect score: 165

Sequence: 1 MKFNKKRAVAIAFIALIFVSPFTISSQDAQAARS 35

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents AA:*

2: /cgn2_6/prodata/1/aa/5A_COMB.pep:*

3: /cgn2_6/prodata/1/aa/6A_COMB.pep:*

4: /cgn2_6/prodata/1/aa/6B_COMB.pep:*

5: /cgn2_6/prodata/1/aa/PCTUS_COMB.pep:*

6: /cgn2_6/prodata/1/aa/backfile1.pep:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	54.5	33.0	351	3	US-09-134-001C-4876	Sequence 4876, Ap
2	52.5	31.8	257	4	US-09-710-279-3244	Sequence 3244, Ap
3	52.5	31.8	264	3	US-09-134-001C-5035	Sequence 5035, Ap
4	52.5	31.8	267	3	US-09-134-001C-4539	Sequence 4539, Ap
5	52	31.5	933	3	US-08-293-728-2	Sequence 2, Appl
6	52	31.5	933	3	US-09-421-868-2	Sequence 2, Appl
7	52	31.5	936	4	US-08-956-171E-5249	Sequence 5249, Ap
8	52	31.5	936	4	US-08-781-986A-5249	Sequence 5249, Ap
9	51.5	31.2	509	3	US-09-198-955A-6	Sequence 6, Appl
10	51.5	31.2	509	3	US-09-694-531-6	Sequence 6, Appl
11	51.5	31.2	509	4	US-10-072-152-6	Sequence 6, Appl
12	50	30.3	418	4	US-09-543-681A-7812	Sequence 7812, Ap
13	50	30.3	1058	2	US-08-247-904B-14	Sequence 14, Appl
14	50	30.3	1058	4	US-08-767-942A-15	Sequence 15, Appl
15	50	30.3	1058	4	US-09-538-092-985	Sequence 985, Ap
16	49	29.7	736	4	US-09-583-110-4555	Sequence 4555, Ap
17	48	29.1	507	4	US-09-248-796A-18283	Sequence 18283, A
18	48	29.1	1391	4	US-09-106-568B-8	Sequence 8, Appl
19	47.5	28.8	409	3	US-09-724-864-53	Sequence 53, Appl
20	47	28.5	363	4	US-09-792-024-80	Sequence 80, Appl
21	47	28.5	2161	1	US-07-745-206A-2	Sequence 2, Appl
22	47	28.5	2161	1	US-08-455-543A-49	Sequence 49, Appl
23	47	28.5	2161	1	US-08-455-543A-51	Sequence 51, Appl
24	47	28.5	2161	2	US-08-223-305C-49	Sequence 49, Appl
25	47	28.5	2161	2	US-08-223-305C-51	Sequence 51, Appl
26	47	28.5	2161	2	US-08-311-363-2	Sequence 2, Appl
27	46.5	28.2	554	3	US-08-591-468-7	Sequence 7, Appl

ALIGNMENTS

28	46.5	28.2	554	4	US-09-650-324A-60	Sequence 60, Appl
29	46.5	28.2	554	5	PCT-US94-06430-7	Sequence 7, Appl
30	46	27.9	111	4	US-09-583-110-5084	Sequence 5084, Ap
31	46	27.9	349	1	US-08-118-270-71	Sequence 71, Appl
32	46	27.9	349	5	PCT-US93-08528-71	Sequence 71, Appl
33	45	27.3	156	4	US-09-328-352-4530	Sequence 4530, Ap
34	45	27.3	180	3	US-09-134-001C-4421	Sequence 4421, Ap
35	45	27.3	703	4	US-09-134-000C-5072	Sequence 5072, Ap
36	44.5	27.0	132	4	US-09-710-279-1438	Sequence 1438, Ap
37	44.5	27.0	141	4	US-09-252-991A-20212	Sequence 20212, A
38	44.5	27.0	323	4	US-09-248-796A-24078	Sequence 24078, A
39	44.5	27.0	684	4	US-09-710-279-2098	Sequence 2098, Ap
40	44.5	27.0	690	3	US-09-134-001C-4568	Sequence 4568, Ap
41	44	26.7	84	4	US-09-543-681A-7825	Sequence 7825, Ap
42	44	26.7	99	4	US-09-270-767-61316	Sequence 61316, A
43	44	26.7	180	4	US-09-393-634-33	Sequence 33, Appl
44	44	26.7	181	4	US-09-270-767-57662	Sequence 57662, A
45	44	26.7	195	4	US-09-328-352-6718	Sequence 6718, Ap

```
RESULT 1
US-09-134-001C-4876
; Sequence 4876, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4876
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4876

Query Match      33.0%; Score 54.5; DB 3; Length 351;
Best local Similarity 41.7%; Pred. No. 3.8;
Matches 15; Conservative 5; Mismatches 15; Indels 1; Gaps 1;

Cy      1 MKFNKKRAVAIAFIALIFVSPFTISSQDAQAARS 35
Db      3 LMKKEKCIWIIAIIILIFVSPFGSGETALTAAART 38

RESULT 2
US-09-710-279-3244
; Sequence 3244, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBLEY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 3244
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```


CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 5249:
SEQUENCE CHARACTERISTICS:
LENGTH: 936 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5249:
US-08-956-171E-5249

Query Match 31.5%; Score 52; DB 4; Length 936;
Best Local Similarity 32.3%; Pred. No. 27;
Matches 10; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 5 KKRVAIATPIALIFVSFFTSSQDAQAARS 35
DB 22 KKSIGAVSLVGTIGFGLSSKEADASENS 52

RESULT 8
US-08-781-986A-5249
Sequence 5249, Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:
APPLICANT: Charles Kunesh
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 5249:
SEQUENCE CHARACTERISTICS:
LENGTH: 936 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-781-986A-5249

Query Match 31.5%; Score 52; DB 4; Length 936;
Best Local Similarity 32.3%; Pred. No. 27;
Matches 10; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 5 KKRVAIATPIALIFVSFFTSSQDAQAARS 35
DB 22 KKSIGAVSLVGTIGFGLSSKEADASENS 52

RESULT 9
US-09-198-955A-6
Sequence 6, Application US/09198955A
Patent No. 6187580
GENERAL INFORMATION:
APPLICANT: Andersen, Lene N.
APPLICANT: Schuelein, Martin
APPLICANT: Lange, Niels E.
APPLICANT: Bjornvad, Made E.
APPLICANT: Moller, Soren
APPLICANT: Glad, Sanne O. S.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Schnorr, Kirk
APPLICANT: Kongsbak, Lars
TITLE OF INVENTION: No. 6187580e1 Pectate Lyases
FILE REFERENCE: 5378, 200-US
CURRENT APPLICATION NUMBER: US/09/198,955A
CURRENT FILING DATE: 1998-11-24
PRIOR APPLICATION NUMBER: 1343/97
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 1344/97
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/067,249
PRIOR FILING DATE: 1997-12-02
PRIOR APPLICATION NUMBER: 60/067,240
PRIOR FILING DATE: 1997-12-02
PRIOR APPLICATION NUMBER: 09/073,684
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 09/184,217
PRIOR FILING DATE: 1998-11-02
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 509
TYPE: prt
ORGANISM: Bacillus sp.
US-09-198-955A-6

Query Match 31.2%; Score 51.5; DB 3; Length 509;
Best Local Similarity 41.4%; Pred. No. 16;
Matches 12; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 5 KKRVAIATPIALIFVSFFTSSQDAQAARS 32
DB 3 KKRKALSLVIVGLFVSFFSFGHGAEAA 31

```
RESULT 10
US-09-694-531-6
; Sequence 6, Application US/09694531
; Patent No. 6368843
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulteis, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Same O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schmoor, Kirk
; APPLICANT: Kongsbak, Lars
; TITLE OF INVENTION: No. 6368843el Pectate Lyases
; FILE REFERENCE: 5378-200-US
; CURRENT APPLICATION NUMBER: US/09/694,531
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 09/198,955
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1343/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/067,249
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 09/184,217
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-694-531-6

Query Match      31.2% Score 51.5; DB 3; Length 509;
Best Local Similarity 41.4%; Pred. No. 16;
Matches 12; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY      5 KKRVAATFIAL-ITVSFTTSSQDAQAA 32
Db      3 KMRKALSVLIVFGLFVSPFSFGHGAAEA 31

RESULT 11
US-10-072-152-6
; Sequence 6, Application US/10072152
; Patent No. 6677147
; GENERAL INFORMATION:
; APPLICANT: Andersen, Lene N.
; APPLICANT: Schulteis, Martin
; APPLICANT: Lange, Niels E.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Moller, Soren
; APPLICANT: Glad, Same O. S.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Schmoor, Kirk
; APPLICANT: Kongsbak, Lars
; TITLE OF INVENTION: No. 6677147el Pectate Lyases
; FILE REFERENCE: 5378-200-US
; CURRENT APPLICATION NUMBER: US/10/072,152
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US/09/198,955
; PRIOR FILING DATE: 1998-11-24
; PRIOR APPLICATION NUMBER: 1343/97
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 1344/97
; PRIOR FILING DATE: 1997-11-24
```

```
; PRIOR APPLICATION NUMBER: 60/067,249
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 60/067,240
; PRIOR FILING DATE: 1997-12-02
; PRIOR APPLICATION NUMBER: 09/073,684
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 09/184,217
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-10-072-152-6

Query Match      31.2% Score 51.5; DB 4; Length 509;
Best Local Similarity 41.4%; Pred. No. 16;
Matches 12; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY      5 KKRVAATFIAL-ITVSFTTSSQDAQAA 32
Db      3 KMRKALSVLIVFGLFVSPFSFGHGAAEA 31

RESULT 12
US-09-543-681A-7812
; Sequence 7812, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7812
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7812

Query Match      30.3% Score 50; DB 4; Length 418;
Best Local Similarity 64.7%; Pred. No. 21;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      8 VAIATFIALITVSPFTI 24
Db      295 IAIIVTIAISIVPFXI 311

RESULT 13
US-08-247-904B-14
; Sequence 14, Application US/08247904B
; Patent No. 5981699
; GENERAL INFORMATION:
; APPLICANT: Rolfe, Mark
; APPLICANT: Eckstein, Jens W.
; APPLICANT: Draetta, Giulio
; TITLE OF INVENTION: Human Ubiquitin Conjugating Enzyme
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Foley, Hong & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/247,904B
FILING DATE: 23-MAY-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIV-029.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1058 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-247-904B-14

Query Match
Best Local Similarity 30.3%; Score 50; DB 2; Length 1058;
Best Local Similarity 32.5%; Pred. No. 60;
Matches 13; Conservative 8; Mismatches 11; Indels 0; Gaps 1;

Qy 4 NKKRAVAITFIALFVSFFT-----ISSQDAQAERS 35
Db 781 SODRAVAITFLOSQVPEFTPKSGVKIHVSDDELQSANMS 820

* RESULT 14
US-08-767-942A-15
Sequence 15, Application US/08/67942A
Patent No. 6068982
GENERAL INFORMATION:
APPLICANT: Rolfe, Mark
APPLICANT: Chiu, M. Isabel
APPLICANT: Berlin, Vivian
APPLICANT: Damagnez, Veronique
APPLICANT: Draela, Giulio
APPLICANT: Guillaume, Cottarel
TITLE OF INVENTION: UBIQUITIN CONJUGATING ENZYMES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESS: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/767,942A
FILING DATE: 17-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIV-029.04
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1058 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-767-942A-15

```

```

Query Match
Best Local Similarity 30.3%; Score 50; DB 3; Length 1058;
Best Local Similarity 32.5%; Pred. No. 60;
Matches 13; Conservative 8; Mismatches 11; Indels 8; Gaps 1;

Qy 4 NKKRAVAITFIALFVSFFT-----ISSQDAQAERS 35
Db 781 SODRAVAITFLOSQVPEFTPKSGVKIHVSDDELQSANMS 820

* RESULT 15
US-09-538-092-985
Sequence 985, Application US/09/538092
Patent No. 6753314
GENERAL INFORMATION:
APPLICANT: Glot, Loic
APPLICANT: Mansfield, Traci A.
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
FILE REFERENCE: 15966-542
CURRENT APPLICATION NUMBER: US/09/538,092
CURRENT FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/127,352
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/178,965
PRIOR FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 1387
SOFTWARE: CuraPatSeqFormatter Version 0.9
SEQ ID NO 985
LENGTH: 1058
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (0)...(0)
OTHER INFORMATION: Polypeptide Accession Number P22314
US-09-538-092-985

Query Match
Best Local Similarity 30.3%; Score 50; DB 4; Length 1058;
Best Local Similarity 32.5%; Pred. No. 60;
Matches 13; Conservative 8; Mismatches 11; Indels 8; Gaps 1;

Qy 4 NKKRAVAITFIALFVSFFT-----ISSQDAQAERS 35
Db 781 SODRAVAITFLOSQVPEFTPKSGVKIHVSDDELQSANMS 820

Search completed: November 7, 2004, 17:57:31
Job time : 26 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 7, 2004, 14:58:20 ; Search time 80 Seconds
(without alignments)
154,536 Million cell updates/sec

Title: US-09-982-531-1

Perfect score: 165
Sequence: 1 MKFNKKRVAIATFIALIFVSPFTISSQDAQAERS 35

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	165	100.0	35	9	US-09-982-531-1
2	157	95.2	35	15	US-10-431-577-48
3	154.5	93.6	36	15	US-10-431-577-43
4	151	91.5	35	15	US-10-431-577-49
5	148.5	90.0	36	15	US-10-431-577-51
6	148.5	90.0	36	15	US-10-431-577-62
7	147.5	89.4	36	15	US-10-431-577-50
8	142.5	86.4	36	15	US-10-431-577-47
9	142.5	86.4	36	15	US-10-431-577-52
10	142	86.1	35	15	US-10-431-577-44
11	141	85.5	34	15	US-10-431-577-59
12	139	84.2	35	15	US-10-431-577-46
13	138.5	83.9	35	15	US-10-431-577-58

14	138.5	83.9	38	15	US-10-431-577-42	Sequence 42, Appl
15	138	83.6	35	15	US-10-431-577-60	Sequence 60, Appl
16	137.5	83.3	38	15	US-10-431-577-41	Sequence 41, Appl
17	135	81.8	35	15	US-10-431-577-45	Sequence 45, Appl
18	126.5	76.7	35	15	US-10-431-577-61	Sequence 61, Appl
19	120.5	73.0	38	15	US-10-431-577-56	Sequence 56, Appl
20	119.5	72.4	38	15	US-10-431-577-53	Sequence 53, Appl
21	105.5	63.9	38	15	US-10-431-577-57	Sequence 57, Appl
22	104.5	63.3	38	15	US-10-431-577-54	Sequence 54, Appl
23	68.5	41.5	38	15	US-10-431-577-55	Sequence 55, Appl
24	58	35.2	273	15	US-10-424-559-267637	Sequence 267637, Appl
25	55	33.3	200	9	US-09-764-853-529	Sequence 529, App
26	54	32.7	88	15	US-10-424-559-201948	Sequence 201948, Appl
27	53	32.1	88	15	US-10-424-559-234743	Sequence 234743, Appl
28	53	32.1	619	15	US-10-282-122A-53626	Sequence 53626, A
29	52	31.5	496	14	US-10-311-879-29	Sequence 29, Appl
30	52	31.5	935	15	US-10-282-122A-44326	Sequence 44326, A
31	52	31.5	936	8	US-08-781-986A-5249	Sequence 5249, A
32	52	31.5	936	15	US-10-329-624-5249	Sequence 5249, Ap
33	52	31.5	1021	9	US-09-815-242-5471	Sequence 5471, Ap
34	52	31.5	1021	9	US-09-815-242-12544	Sequence 12544, A
35	51.5	31.2	287	14	US-10-156-761-9477	Sequence 9477, Ap
36	51.5	31.2	509	13	US-10-072-152-6	Sequence 6, Appl
37	51.5	31.2	509	15	US-10-655-433-6	Sequence 6, Appl
38	51.5	31.2	952	16	US-10-437-963-115492	Sequence 115492, Appl
39	50.5	30.6	258	9	US-09-815-242-5894	Sequence 5894, Ap
40	50.5	30.6	690	9	US-09-815-242-5841	Sequence 5841, Ap
41	50.5	30.6	691	15	US-10-282-122A-44369	Sequence 44369, A
42	50	30.3	212	10	US-09-905-666A-55	Sequence 55, Appl
43	50	30.3	212	10	US-09-905-666A-56	Sequence 56, Appl
44	50	30.3	212	10	US-09-905-666A-58	Sequence 58, Appl
45	50	30.3	212	10	US-09-905-666A-59	Sequence 59, Appl

ALIGNMENTS

RESULT 1
US-09-982-531-1
Sequence 1, Application US/09982531
Patent No. US20020137140A1
GENERAL INFORMATION:
APPLICANT: Vang, Astrid
APPLICANT: Madson, Soren
APPLICANT: Bredemose, Lars
APPLICANT: Ravn, Peter
APPLICANT: Arnan, Jose
APPLICANT: Johnsen, Mads
APPLICANT: Steenberg, Anne
APPLICANT: Israelien, Hans
TITLE OF INVENTION: Improved Fermentation Method for Production of Heterologous Gene
FILE REFERENCE: 54320.000011
CURRENT APPLICATION NUMBER: US/09/982,531
CURRENT FILING DATE: 2001-10-19
PRIOR APPLICATION NUMBER: US 09/692,205
PRIOR FILING DATE: 2000-10-20
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 35
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: SP10mut2
US-09-982-531-1
Query Match 100.0%; Score 165; DB 9; Length 35;
Best Local Similarity 100.0%; Pred. No. 9.1e-17;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MKFNKKRVAIATFIALIFVSPFTISSQDAQAERS 35
|||||

Db 1 MKFNKKRVAIAATFIALIFVSPFTTSSQDAQAERS 35

RESULT 2

US-10-431-577-48
; Sequence 48, Application US/10431577
; Publication No. US20040038263A1
; GENERAL INFORMATION:
; APPLICANT: RAVN, PETER
; APPLICANT: MADSEN, SOEREN MICHAEL
; APPLICANT: VRANG, ASTRID
; APPLICANT: ISRAEISEN, HANS
; APPLICANT: MADSEN, MADS GROENVOLD
; APPLICANT: BREDMOSE, LARS
; APPLICANT: ARNAU, JOSE
; TITLE OF INVENTION: METHOD OF ISOLATING SECRETION SIGNALS IN LACTIC ACID
; TITLE OF INVENTION: BACTERIA AND NOVEL SECRETION SIGNALS ISOLATED FROM
; FILE REFERENCE: 54320.000012
; CURRENT APPLICATION NUMBER: US/10/431,577
; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-431-577-48

Query Match 95.2%; Score 157; DB 15; Length 35;
Best Local Similarity 97.1%; Pred. No. 1,3e-15;
Matches 34; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKFNKKRVAIAATFIALIFVSPFTTSSQDAQAERS 35

Db 1 MKFNKKRVAIAATFIALIFVSPFTTSSQDAQAERS 35

RESULT 3

US-10-431-577-43
; Sequence 43, Application US/10431577
; Publication No. US20040038263A1
; GENERAL INFORMATION:
; APPLICANT: RAVN, PETER
; APPLICANT: MADSEN, SOEREN MICHAEL
; APPLICANT: VRANG, ASTRID
; APPLICANT: ISRAEISEN, HANS
; APPLICANT: MADSEN, MADS GROENVOLD
; APPLICANT: BREDMOSE, LARS
; APPLICANT: ARNAU, JOSE
; TITLE OF INVENTION: METHOD OF ISOLATING SECRETION SIGNALS IN LACTIC ACID
; TITLE OF INVENTION: BACTERIA AND NOVEL SECRETION SIGNALS ISOLATED FROM
; FILE REFERENCE: 54320.000012
; CURRENT APPLICATION NUMBER: US/10/431,577
; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-431-577-43

Query Match 93.6%; Score 154.5; DB 15; Length 36;
Best Local Similarity 97.2%; Pred. No. 3,2e-15;
Matches 35; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MKFNKKRVAIAATFIALIFVSPFTTSSQDAQAERS 35

Db 1 MKFNKKRVAIAATFIALIFVSPFTTSSQDAQAERS 36

RESULT 4

US-10-431-577-49
; Sequence 49, Application US/10431577
; Publication No. US20040038263A1
; GENERAL INFORMATION:
; APPLICANT: RAVN, PETER
; APPLICANT: MADSEN, SOEREN MICHAEL
; APPLICANT: VRANG, ASTRID
; APPLICANT: ISRAEISEN, HANS
; APPLICANT: MADSEN, MADS GROENVOLD
; APPLICANT: BREDMOSE, LARS
; APPLICANT: ARNAU, JOSE
; TITLE OF INVENTION: METHOD OF ISOLATING SECRETION SIGNALS IN LACTIC ACID
; TITLE OF INVENTION: BACTERIA AND NOVEL SECRETION SIGNALS ISOLATED FROM
; FILE REFERENCE: 54320.000012
; CURRENT APPLICATION NUMBER: US/10/431,577
; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-431-577-49

Query Match 91.5%; Score 151; DB 15; Length 35;
Best Local Similarity 94.3%; Pred. No. 1e-14;
Matches 33; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKFNKKRVAIAATFIALIFVSPFTTSSQDAQAERS 35

Db 1 MKFNKKRVAIAATFIALIFVSPFTTSSQDAQAERS 35

RESULT 5

US-10-431-577-51
; Sequence 51, Application US/10431577
; Publication No. US20040038263A1
; GENERAL INFORMATION:
; APPLICANT: RAVN, PETER
; APPLICANT: MADSEN, SOEREN MICHAEL
; APPLICANT: VRANG, ASTRID
; APPLICANT: ISRAEISEN, HANS
; APPLICANT: MADSEN, MADS GROENVOLD
; APPLICANT: BREDMOSE, LARS
; APPLICANT: ARNAU, JOSE
; TITLE OF INVENTION: METHOD OF ISOLATING SECRETION SIGNALS IN LACTIC ACID
; TITLE OF INVENTION: BACTERIA AND NOVEL SECRETION SIGNALS ISOLATED FROM
; FILE REFERENCE: 54320.000012
; CURRENT APPLICATION NUMBER: US/10/431,577
; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-431-577-51

Query Match 93.6%; Score 154.5; DB 15; Length 36;
Best Local Similarity 97.2%; Pred. No. 3,2e-15;
Matches 35; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Query Match 90.0%; Score 148.5; DB 15; Length 36;
Best Local Similarity 94.4%; Pred. No. 2,4e-14;
Matches 34; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 MKFNKRVAIATFIALIFVSFFTIS-ODAOAERS 35
1 MKFNKRVAIATFIALIFVSFFTISIDAOAERS 36

RESULT 6
US-10-431-577-62

Sequence 62, Application US/10431577
Publication No. US20040038263A1

GENERAL INFORMATION:

APPLICANT: RAVN, PETER

APPLICANT: MADSEN, SOEREN MICHAEL

APPLICANT: VRANG, ASTRID

APPLICANT: ISRAELSEN, HANS

APPLICANT: JOHNSEN, MADS GROENVOLD

APPLICANT: BREDMOSE, LARS

APPLICANT: ARNAU, JOSE

TITLE OF INVENTION: METHOD OF ISOLATING SECRETION SIGNALS IN LACTIC ACID

TITLE OF INVENTION: BACTERIA AND NOVEL SECRETION SIGNALS ISOLATED FROM

FILE REFERENCE: 54320.000012

CURRENT APPLICATION NUMBER: US/10/431,577

CURRENT FILING DATE: 2003-05-08

NUMBER OF SEQ ID NOS: 66

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 62

LENGTH: 36

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

OTHER INFORMATION: amino acid sequence

US-10-431-577-62

Query Match 90.0%; Score 148.5; DB 15; Length 36;
Best Local Similarity 94.4%; Pred. No. 2,4e-14;
Matches 34; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 MKFNKRVAIATFIALIFVSFFTIS-ODAOAERS 35
1 MKFNKRVAIATFIALIFVSFFTISIDAOAERS 36

RESULT 7
US-10-431-577-50

Sequence 50, Application US/10431577
Publication No. US20040038263A1

GENERAL INFORMATION:

APPLICANT: RAVN, PETER

APPLICANT: MADSEN, SOEREN MICHAEL

APPLICANT: VRANG, ASTRID

APPLICANT: ISRAELSEN, HANS

APPLICANT: JOHNSEN, MADS GROENVOLD

APPLICANT: BREDMOSE, LARS

APPLICANT: ARNAU, JOSE

TITLE OF INVENTION: METHOD OF ISOLATING SECRETION SIGNALS IN LACTIC ACID

TITLE OF INVENTION: BACTERIA AND NOVEL SECRETION SIGNALS ISOLATED FROM

FILE REFERENCE: 54320.000012

CURRENT APPLICATION NUMBER: US/10/431,577

CURRENT FILING DATE: 2003-05-08

NUMBER OF SEQ ID NOS: 66

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 50

LENGTH: 36

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

OTHER INFORMATION: amino acid sequence
US-10-431-577-50

Query Match 89.4%; Score 147.5; DB 15; Length 36;
Best Local Similarity 94.4%; Pred. No. 3,3e-14;
Matches 34; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 MKFNKRVAIATFIALIFVSFFTIS-ODAOAERS 35
1 MKFNKRVAIATFIALIFVSFFTISIDAOAERS 36

RESULT 8
US-10-431-577-47

Sequence 47, Application US/10431577
Publication No. US20040038263A1

GENERAL INFORMATION:

APPLICANT: RAVN, PETER

APPLICANT: MADSEN, SOEREN MICHAEL

APPLICANT: VRANG, ASTRID

APPLICANT: ISRAELSEN, HANS

APPLICANT: JOHNSEN, MADS GROENVOLD

APPLICANT: BREDMOSE, LARS

APPLICANT: ARNAU, JOSE

TITLE OF INVENTION: METHOD OF ISOLATING SECRETION SIGNALS IN LACTIC ACID

TITLE OF INVENTION: BACTERIA AND NOVEL SECRETION SIGNALS ISOLATED FROM

FILE REFERENCE: 54320.000012

CURRENT APPLICATION NUMBER: US/10/431,577

CURRENT FILING DATE: 2003-05-08

NUMBER OF SEQ ID NOS: 66

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 47

LENGTH: 36

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

OTHER INFORMATION: amino acid sequence

US-10-431-577-47

Query Match 86.4%; Score 142.5; DB 15; Length 36;
Best Local Similarity 91.7%; Pred. No. 1,8e-13;
Matches 33; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 MKFNKRVAIATFIALIFVSFFTIS-ODAOAERS 35
1 MKFNKRVAIATFIALIFVSFFTISIDAOAERS 36

RESULT 9
US-10-431-577-52

Sequence 52, Application US/10431577
Publication No. US20040038263A1

GENERAL INFORMATION:

APPLICANT: RAVN, PETER

APPLICANT: MADSEN, SOEREN MICHAEL

APPLICANT: VRANG, ASTRID

APPLICANT: ISRAELSEN, HANS

APPLICANT: JOHNSEN, MADS GROENVOLD

APPLICANT: BREDMOSE, LARS

APPLICANT: ARNAU, JOSE

TITLE OF INVENTION: METHOD OF ISOLATING SECRETION SIGNALS IN LACTIC ACID

TITLE OF INVENTION: BACTERIA AND NOVEL SECRETION SIGNALS ISOLATED FROM

FILE REFERENCE: 54320.000012

CURRENT APPLICATION NUMBER: US/10/431,577

CURRENT FILING DATE: 2003-05-08

NUMBER OF SEQ ID NOS: 66

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 52

LENGTH: 36

TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: amino acid sequence
US-10-431-577-52

Query Match 86.4%; Score 142.5; DB 15; Length 36;
Best Local Similarity 91.7%; Pred. No. 1.8e-13;
Matches 33; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 MKFNKRVATFIALIFVSFFTSSODQAARS 35
DB 1 MKFNKRVATFIALIFVSFFTSSODQAARS 36

RESULT 10

US-10-431-577-44
Sequence 44, Application US/10431577
Publication No. US20040038263A1
GENERAL INFORMATION:
APPLICANT: RAVN, PETER
APPLICANT: MADSEN, SOEREN MICHAEL
APPLICANT: VRANG, ASTRID
APPLICANT: ISRAEISEN, HANS
APPLICANT: JOHNSEN, MADS GROENVOLD
APPLICANT: BREDMOSE, LARS
APPLICANT: ARNAU, JOSE
TITLE OF INVENTION: METHOD OF ISOLATING SECRETION SIGNALS IN LACTIC ACID
TITLE OF INVENTION: BACTERIA AND NOVEL SECRETION SIGNALS ISOLATED FROM
TITLE OF INVENTION: LACTOCOCCUS LACTIS
FILE REFERENCE: 54320.000012
CURRENT APPLICATION NUMBER: US/10/431,577
CURRENT FILING DATE: 2003-05-08
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 44
LENGTH: 35
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: amino acid sequence
US-10-431-577-44

Query Match 86.1%; Score 142; DB 15; Length 35;
Best Local Similarity 88.6%; Pred. No. 2e-13;
Matches 31; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKFNKRVATFIALIFVSFFTSSODQAARS 35
DB 1 MKFNKRVATFIALIFVSFFTIIPTQAARS 35

RESULT 11

US-10-431-577-59
Sequence 59, Application US/10431577
Publication No. US20040038263A1
GENERAL INFORMATION:
APPLICANT: RAVN, PETER
APPLICANT: MADSEN, SOEREN MICHAEL
APPLICANT: VRANG, ASTRID
APPLICANT: ISRAEISEN, HANS
APPLICANT: JOHNSEN, MADS GROENVOLD
APPLICANT: BREDMOSE, LARS
APPLICANT: ARNAU, JOSE
TITLE OF INVENTION: METHOD OF ISOLATING SECRETION SIGNALS IN LACTIC ACID
TITLE OF INVENTION: BACTERIA AND NOVEL SECRETION SIGNALS ISOLATED FROM
TITLE OF INVENTION: LACTOCOCCUS LACTIS
FILE REFERENCE: 54320.000012
CURRENT APPLICATION NUMBER: US/10/431,577
CURRENT FILING DATE: 2003-05-08
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 59
LENGTH: 34
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: amino acid sequence
US-10-431-577-59

Query Match 85.5%; Score 141; DB 15; Length 34;
Best Local Similarity 96.9%; Pred. No. 2.8e-13;
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 NKRVATFIALIFVSFFTSSODQAARS 35
DB 3 NKRVATFIALIFVSFFTSSODQAARS 34

RESULT 12

US-10-431-577-46
Sequence 46, Application US/10431577
Publication No. US20040038263A1
GENERAL INFORMATION:
APPLICANT: RAVN, PETER
APPLICANT: MADSEN, SOEREN MICHAEL
APPLICANT: VRANG, ASTRID
APPLICANT: ISRAEISEN, HANS
APPLICANT: JOHNSEN, MADS GROENVOLD
APPLICANT: BREDMOSE, LARS
APPLICANT: ARNAU, JOSE
TITLE OF INVENTION: METHOD OF ISOLATING SECRETION SIGNALS IN LACTIC ACID
TITLE OF INVENTION: BACTERIA AND NOVEL SECRETION SIGNALS ISOLATED FROM
TITLE OF INVENTION: LACTOCOCCUS LACTIS
FILE REFERENCE: 54320.000012
CURRENT APPLICATION NUMBER: US/10/431,577
CURRENT FILING DATE: 2003-05-08
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 46
LENGTH: 35
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: amino acid sequence
US-10-431-577-46

Query Match 84.2%; Score 139; DB 15; Length 35;
Best Local Similarity 85.7%; Pred. No. 5.6e-13;
Matches 30; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKFNKRVATFIALIFVSFFTSSODQAARS 35
DB 1 MKFNKRVATFIALIFVSFTIIQNAARS 35

RESULT 13

US-10-431-577-58
Sequence 58, Application US/10431577
Publication No. US20040038263A1
GENERAL INFORMATION:
APPLICANT: RAVN, PETER
APPLICANT: MADSEN, SOEREN MICHAEL
APPLICANT: VRANG, ASTRID
APPLICANT: ISRAEISEN, HANS
APPLICANT: JOHNSEN, MADS GROENVOLD
APPLICANT: BREDMOSE, LARS
APPLICANT: ARNAU, JOSE
TITLE OF INVENTION: METHOD OF ISOLATING SECRETION SIGNALS IN LACTIC ACID
TITLE OF INVENTION: BACTERIA AND NOVEL SECRETION SIGNALS ISOLATED FROM
TITLE OF INVENTION: LACTOCOCCUS LACTIS
FILE REFERENCE: 54320.000012
CURRENT APPLICATION NUMBER: US/10/431,577

```

; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-431-577-58

Query Match      83.9%; Score 138.5; DB 15; Length 35;
Best Local Similarity 97.0%; Pred. No. 6, 6e-13;
Matches 32; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy      4 NKKRVAIATFIALIFVSPFTISS-IDQAAERS 35
Db      3 NKKRVAIATFIALIFVSPFTISS-IDQAAERS 35

RESULT 14
US-10-431-577-42
; Sequence 42, Application US/10431577
; Publication No. US20040038263A1
; GENERAL INFORMATION:
; APPLICANT: RAVN, PETER
; APPLICANT: MADSEN, SOEREN MICHAEL
; APPLICANT: VRANG, ASTRID
; APPLICANT: ISRAELSEN, HANS
; APPLICANT: JOHNSEN, MADS GROENVOLD
; APPLICANT: BREDMOSE, LARS
; APPLICANT: ARNAV, JOSE
; TITLE OF INVENTION: METHOD OF ISOLATING SECRETION SIGNALS IN LACTIC ACID
; TITLE OF INVENTION: BACTERIA AND NOVEL SECRETION SIGNALS ISOLATED FROM
; FILE REFERENCE: 54320.000012
; CURRENT APPLICATION NUMBER: US/10/431,577
; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-431-577-42

Query Match      83.9%; Score 138.5; DB 15; Length 36;
Best Local Similarity 89.5%; Pred. No. 7, 2e-13;
Matches 34; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

Qy      1 MKENKRVAIATFIALIFVSPFTISS-OD--AQAERS 35
Db      1 MKENKRVAIATFIALIFVSPFTISS-IDQAAERS 38

RESULT 15
US-10-431-577-60
; Sequence 60, Application US/10431577
; Publication No. US20040038263A1
; GENERAL INFORMATION:
; APPLICANT: RAVN, PETER
; APPLICANT: MADSEN, SOEREN MICHAEL
; APPLICANT: VRANG, ASTRID
; APPLICANT: ISRAELSEN, HANS
; APPLICANT: JOHNSEN, MADS GROENVOLD
; APPLICANT: BREDMOSE, LARS
; APPLICANT: ARNAV, JOSE
; TITLE OF INVENTION: METHOD OF ISOLATING SECRETION SIGNALS IN LACTIC ACID
; TITLE OF INVENTION: BACTERIA AND NOVEL SECRETION SIGNALS ISOLATED FROM
```

```

; TITLE OF INVENTION: LACTOCOCCUS LACTIS
; FILE REFERENCE: 54320.000012
; CURRENT APPLICATION NUMBER: US/10/431,577
; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-431-577-60

Query Match      83.6%; Score 138; DB 15; Length 35;
Best Local Similarity 94.4%; Pred. No. 7, 8e-13;
Matches 34; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy      1 MKENKRVAIATFIALIFVSPFTISS-IDQAAERS 35
Db      1 MKE-KRVAIATFIALIFVSPFTISS-IDQAAERS 35
```

Search completed: November 7, 2004, 17:58:58
Job time : 81 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2004, 14:58:19 ; Search time 24 Seconds
(without alignments)
140.316 Million cell updates/sec

Title: US-09-982-531-1

Perfect score: 165
Sequence: 1 MKFNKKRVAIAATFIALIFVSFPTISSQDAQAALERS 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

* Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56.5	34.2	411	2	A37755 xylanase (EC 3.2.1.1)
2	53.5	32.4	447	2	B82490 probable Na+/H+ an
3	53	32.1	410	2	F88978 protein F37B4.4 [1
4	53	32.1	550	1	A26076 UDP-sugar diphosph
5	52.5	31.8	521	2	AC1525 internalin like pr
6	52	31.5	141	2	G90242 signal peptidase,
7	52	31.5	933	2	S41539 fibrinogen-binding
8	52	31.5	989	2	D89852 fibritogen-binding
9	51	30.9	132	2	G70314 conserved hypochet
10	51	30.9	701	2	S48452 probable membrane
11	50.5	30.6	690	2	A24545 triacylglycerol 11
12	50.5	30.6	691	2	B89797 glycerol ester hyd
13	50	30.3	212	2	S23934 lipase 11pA - Baci
14	50	30.3	306	2	C82242 oligopeptide ABC t
15	50	30.3	349	2	D70141 oligopeptide ABC t
16	50	30.3	1058	2	A38564 ubiquitin-protein
17	49.5	30.0	404	2	H97055 membrane bound tra
18	49.5	30.0	750	2	A90590 hypothetical prote
19	49	29.7	111	2	A95102 hypothetical prote
20	49	29.7	111	2	B97970 conserved hypochet
21	49	29.7	383	2	H90266 D-alanyl-D-alanine
22	49	29.7	445	2	AC1794 transporter BMR102
23	49	29.7	481	2	AC3281 prolyl oligopeptid
24	49	29.7	705	2	UK0194 hypothetical prote
25	49	29.7	1179	2	E98012 pili protein (clon
26	48.5	29.4	83	2	T03673 probable multizid
27	48.5	29.4	393	2	AG0184 hypothetical prote
28	48	29.1	131	2	B69831 hypothetical prote
29	48	29.1	345	2	T24533 hypothetical prote

30	48	29.1	368	2	F81414 probable transmembr
31	48	29.1	445	2	A11418 D-alanyl-D-alanine
32	48	29.1	586	2	AD2493 hypothetical prote
33	48	29.1	956	2	A57121 thrombospondin 3 p
34	48	29.1	1062	2	D96540 hypothetical prote
35	47.5	28.8	147	2	C82925 hypothetical prote
36	47.5	28.8	399	2	D70072 antibiotic resistanc
37	47.5	28.8	590	2	T20153 hypothetical prote
38	47.5	28.8	736	2	D90574 hypothetical prote
39	47.5	28.8	906	2	B96901 uncharacterized co
40	47	28.5	67	2	F90057 hypothetical prote
41	47	28.5	284	2	B84118 stage II sporulati
42	47	28.5	301	2	S50737 probable membrane
43	47	28.5	302	2	A99847 oligopeptide trans
44	47	28.5	302	2	H85704 oligopeptide trans
45	47	28.5	302	2	H64871 oligopeptide trans

ALIGNMENTS

RESULT 1

A37755

xylanase (EC 3.2.1.-) precursor - Butyrivibrio fibrisolvans

C:Species: Butyrivibrio fibrisolvans

C>Date: 19-Apr-1991 #sequence_revision 19-Apr-1991 #ext_change 09-Jul-2004

C:Accession: A37755

R:Mannarelli, B.M.; Evans, S.; Lee, D.

J. Bacteriol. 172, 4247-4254, 1990

A:Title: Cloning, sequencing, and expression of a xylanase gene from the anaerobic rumine

A:Reference number: A37755; MUID:90330526; PMID:2198249

A:Accession: A37755

A:Status: preliminary; not compared with conceptual translation

A:Residues: 1-411 <MAN>

A:Molecule type: DNA

A:Cross-references: UNIPROT:P2351

C:Superfamily: Bacillus endo-beta-1,4-xylanase; Streptomyces endo-1,4-beta-xylanase A hom

C:Keywords: glycosidase; hydrolase

F:66-383/Domain: Streptomyces endo-1,4-beta-xylanase A homology <SXY>

Query Match

Best Local Similarity 43.8%; Pred. No. 1.8; Length 411;

Matches 14; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

Qy

2 KFN-KKRVAAIAATFIALIFVSFPTISSQDAQA 32

Db 3 KFKIRKLMAVLAALVSTFPMVSKVDANNA 34

RESULT 2

B82490 probable Na+/H+ antiporter VCA0193 [imported] - Vibrio cholerae (strain N16961 serogroup

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #ext_change 09-Jul-2004

C:Accession: B82490

R:Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.V.;

Charadon, D.; Ermolaeva, M.D.; Yamahewyan, J.J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: B82490

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-447 <HEI>

A:Cross-references: UNIPROT:Q9KMX3; GB:AE004359; GB:AE003853; NID:G9657575; PIDN:AAF96106

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VCA0193

A:Map position: 2

C:Superfamily: Na+/H+-exchanging protein

Query Match

32.4%; Score 53.5; DB 2; Length 447;

Best Local Similarity 44.1%; Pred. No. 5.3;
Matches 15; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

Qy 2 KENKRAVAITFALIFVSPFTISSQDAQAERS 35
Db 200 KF-KENIRIALPALALVFTFSSQTAQSPOTS 232

RESULT 3

P88978

Protein F37B4.4 [imported] - *Caenorhabditis elegans*C/Species: *Caenorhabditis elegans*

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C/Accession: F88978

R/anonymous, The C. elegans Sequencing Consortium.

A/Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology

A/Reference number: A75000; MUID:99069613; PMID:9851916

A/Note: see webstiles genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_ele

A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A/Accession: F88978

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1410 <STO>

A/Cross-references: GB:chr_V; PIDN:AA04408.1; PID:g2911837; GSPDB:GN00023; CESP:F37B4.4

C/Genetics:

A/Gene: F37B4.4

A/Map position: 5

Query Match 32.1%; Score 53; DB 2; Length 410;
Best Local Similarity 40.7%; Pred. No. 5.9;
Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 8 VAATFATLIFVSPFTISSQDAQAER 34
Db 361 IVVASFLVPIPLTFSNMGDAWASR 387

RESULT 4

A26076

UDP-sugar diphosphatase (EC 3.6.1.45) precursor - *Salmonella typhimurium*

N/Alternate names: UDPglucose hydrolase

N/Contains: 5'-nucleotide (EC 3.1.3.5)

C/Species: *Salmonella typhimurium*

C/Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004

C/Accession: A26076

R/Burns, D.M.; Beacham, I.R.

J. Mol. Biol. 192, 163-175, 1986

A/Title: Identification and sequence analysis of a silent gene (ushA(0)) in *Salmonella ty*

A/Reference number: A26076; MUID:87169716; PMID:3031310

A/Accession: A26076

A/Molecule type: DNA

A/Residues: 1-550 <BUN>

A/Cross-references: UNIPROT:P06196; GB:X04651; NID:947949; PIDN:CAA28348.1; PID:g47950

C/Comment: This protein, though functional, is not detectably expressed in this organism

C/Genetics:

A/Gene: ushA(0)

C/Superfamily: 5'-nucleotidase; 5'-nucleotidase homology; phosphoesterase core homology

C/Keywords: phosphoric monoester hydrolase

F1-22/Domain: signal sequence #status predicted <SIG>

F123-550/Product: UDP-sugar hydrolase #status predicted <MAT>

F132-526/Domain: 5'-nucleotidase homology <NMS>

F135-118/Domain: phosphoesterase core homology <PEC>

Query Match 32.1%; Score 53; DB 1; Length 550;
Best Local Similarity 41.2%; Pred. No. 7.6;
Matches 14; Conservative 6; Mismatches 8; Indels 6; Gaps 1;

Qy 1 MKFNKRAVAITFALIFVSPFTISSQDAQAER 34
Db 1 MKFLKRGVALALAA-----FALTTPADAYEK 28

RESULT 5
AC1525
internalin like protein (LpxXG) [imported] - *Listeria innocua* (strain C1p11262)

C/Species: *Listeria innocua*

C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C/Accession: AC1525

R/Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,

D.; Dominguez-Bernal, G.; Duhaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Feihl, H.;

Science 294, 849-852, 2001

A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madeno, E.; Maicournam, A.; Mat

Ok, C.; Schueter, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Vose, H.; Wehlend,

A/Title: Comparative genomics of *Listeria* species.

A/Reference number: AB1077; MUID:21537279; PMID:11679669

A/Accession: AC1525

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-521 <GUA>

A/Cross-references: UNIPROT:Q92D53; GB:AL592022; PIDN:CAC95971.1; PID:G16413191; GSPDB:G

A/Experimental source: strain C1p11262

C/Genetics:

A/Gene: lin0739

Query Match 31.8%; Score 52.5; DB 2; Length 521;
Best Local Similarity 36.1%; Pred. No. 8.6;
Matches 13; Conservative 11; Mismatches 9; Indels 3; Gaps 2;

Qy 1 MKFNKRAVAITFALIFVSPFTISSQDAQAERS 35
Db 2 INWKKITWISLVSLVLSF--ISSPNAQAERS 35

RESULT 6

G90242

signal peptidase, probable [imported] - *Sulfolobus solfataricus*C/Species: *Sulfolobus solfataricus*

C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C/Accession: G90242

R/She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyaz, M.J.; Chan-Y

arct, I.; Jettles, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.

submitted to Genbank, April 2001

A/Description: *Sulfolobus solfataricus* complete genome.

A/Reference number: A99139

A/Accession: G90242

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-141 <KUR>

A/Cross-references: UNIPROT:Q97ZJ0; GB:AE006641; NID:G13814098; PIDN:AAK41198.1; GSPDB:G

C/Genetics:

A/Gene: SSO0916

Query Match 31.5%; Score 52; DB 2; Length 141;
Best Local Similarity 41.4%; Pred. No. 3.2;
Matches 12; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

Qy 1 MKFNKRAVAITFALIFVSPFTISSQDA 29
Db 1 MKMKSDIIITLITLITLMSNIVQSA 29

RESULT 7

S41539

fibrinogen-binding protein - *Staphylococcus aureus*

N/Alternate names: clumping factor

C/Species: *Staphylococcus aureus*

C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C/Accession: S41539; S36630

R/McDevitt, D.; Francois, P.; Vaudaux, P.; Foster, T.J.

Mol. Microbiol. 11, 237-248, 1994

A/Title: Molecular characterization of the clumping factor (fibrinogen receptor) of *Staph*

A/Reference number: S41539; MUID:94224142; PMID:8170386

A/Accession: S41539

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-933 <MCD>
A/Cross-references: UNIPROT:Q53653; EMBL:Z18852; NID:G939752; PIDN:CAA79304.1; PID:G93975

Query Match 31.5%; Score 52; DB 2; Length 933;
Best Local Similarity 32.3%; Pred. No. 17;
Matches 10; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 5 KKRVAITATIALIFVSPFTISSODAAERS 35
DB 13 KKSIGVAVLTGLIGFGLSSKEADASENS 43

RESULT 8

D89852
Elastinogen-binding protein A, clumping factor (imported) - Staphylococcus aureus (strain

C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: D89852

R.Kuroda, M.; Ohba, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Uti, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramoto, K.
Lancet 357, 1225-1240, 2001

A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A89758; NUID:21311952; PMID:11418146

A/Accession: D89852

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-989 <KUR>

A/Cross-references: UNIPROT:Q99VU4; GB:BA000018; PID:G13700678; PIDN:BA81975.1; GSPDB:G

A/Experimental source: strain N315

C/Genetics:

A/Genes: cIfa

Query Match 31.5%; Score 52; DB 2; Length 989;
Best Local Similarity 32.3%; Pred. No. 18;
Matches 10; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 5 KKRVAITATIALIFVSPFTISSODAAERS 35
DB 13 KKSIGVAVLTGLIGFGLSSKEADASENS 43

RESULT 9

G70314
conserved hypothetical protein aq_156 - Aquifex aeolicus

C/Species: Aquifex aeolicus

C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004

C/Accession: G70314

R.Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov

V.
Nature 392, 353-358, 1998

A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A/Reference number: A70300; NUID:98196666; PMID:9537320

A/Accession: G70314

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-132 <AOF>

A/Cross-references: UNIPROT:O66546; GB:AE000676; NID:G2982884; PIDN:AA06515.1; PID:G298

A/Experimental source: strain VFS

C/Genetics:

A/Genes: aq_156

C/Superfamily: Bacillus subtilis hypothetical protein yozB

Query Match 30.9%; Score 51; DB 2; Length 132;
Best Local Similarity 40.0%; Pred. No. 4.2;
Matches 10; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 2 KKKRVAITATIALIFVSPFTISS 26
DB 33 EWHKRAMLTASFALIFVLTLYIKS 57

RESULT 10

S48452
probable membrane protein YII005w - yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein YIA5w

C/Species: Saccharomyces cerevisiae

C/Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004

C/Accession: S48452; S58676

R.Rowley, N.

submitted to the EMBL Data Library, August 1994

A/Reference number: S48442

A/Accession: S48452

A/Molecule type: DNA

A/Residues: 1-701 <ROW>

A/Cross-references: UNIPROT:P40557; GB:Z47047; EMBL:Z38113; NID:G603997; PID:G763341; MII

R.Voss, H.; Tamames, J.; Teodoru, C.; Valencic, A.; Sensen, C.; Wiemann, S.; Schwager, C.

Yeast 11, 61-78, 1995

A/Title: Nucleotide sequence and analysis of the centromeric region of yeast chromosome 1

A/Reference number: S50795; NUID:95282515; PMID:7762303

A/Accession: S58676

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 549-701 <VOS>

A/Cross-references: EMBL:X79743

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994

C/Genetics:

A/Cross-references: SGD:S0001267

A/Map position: 9L

C/Keywords: transmembrane protein

F.8-24/Domain: transmembrane #status predicted <TM1>

F.655-671/Domain: transmembrane #status predicted <TM2>

Query Match 30.9%; Score 51; DB 2; Length 701;
Best Local Similarity 35.7%; Pred. No. 18;
Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 MKENKRVATATIALIFVSPFTISSQD 28
DB 1 MKNMLRLVVTFFSCITFLKFTIAAE 28

RESULT 11

A24545
triacylglycerol lipase (EC 3.1.1.3) - Staphylococcus aureus

C/Species: Staphylococcus aureus

C/Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004

C/Accession: A24545

R.Lee, C.Y.; Iandolo, J.J.

J. Bacteriol. 166, 385-391, 1986

A/Title: Lyso-genic conversion of staphylococcal lipase is caused by insertion of the bact

A/Reference number: A24545; NUID:86195821; PMID:3009394

A/Accession: A24545

A/Molecule type: DNA

A/Residues: 1-690 <LEE>

A/Cross-references: UNIPROT:P10335; GB:M12715; NID:G153019; PIDN:AAA26633.1; PID:G153020

C/Superfamily: Staphylococcus triacylglycerol lipase

C/Keywords: carboxylic ester hydrolase

Query Match 30.6%; Score 50.5; DB 2; Length 690;
Best Local Similarity 26.5%; Pred. No. 21;
Matches 9; Conservative 15; Mismatches 9; Indels 1; Gaps 1;

QY 2 KKKRVAITATIALIFVSPFTISSQDAAERS 35
DB 9 KYSIRKYSIGV-VSVIATATMFVSSHEAASERT 41

RESULT 12
B89797
glycerol ester hydrolase (imported) - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: B89797

R./Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.; C./Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsugu, K.
 Lancel 357, 1255-1240, 2001
 A>Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A/Reference number: A89758; MUID:21311952; PMID:11418146
 A/Accession: B89797
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-691 <KUR>
 A/Cross-references: UNIPROT:Q99W06; GB:BA000018; PID:913700235; PIDN:BA841533.1; GSPDB:G
 A/Experimental source: strain N315
 C/Genetics:
 A/Gene: *geh*
 C/Superfamily: *Staphylococcus triacylglycerol lipase*

Query Match 30.6%; Score 50.5; DB 2; Length 691;
 Best Local Similarity 26.5%; Pred. No. 21;
 Matches 9; Conservative 15; Mismatches 9; Indels 1; Gaps 1;

QY 2 KFNKKRVAATFALFVSFFITSSQDAQAERS 35
 Db 9 KVEIRKSTGV-VSVLAATFVVSHEAQSSEKT 41

RESULT 13
 S23934
 Lipase 11pA - *Bacillus subtilis*
 C/Species: *Bacillus subtilis*
 C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 C/Accession: S23934; B69652
 R/Barlois, V.; Baulard, A.; Schanck, K.; Colson, C.
 Biochim. Biophys. Acta 1131, 253-260, 1992
 A>Title: Cloning, nucleotide sequence and expression in *Escherichia coli* of a lipase gen
 A/Reference number: S23934; MUID:92329538; PMID:1320940
 A/Accession: S23934
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-212 <DAR>

A/Cross-references: UNIPROT:P37957; GB:M74010; NID:9143153; PIDN:AA22574.1; PID:9143154
 R/Kumst, F.; Ogasawara, N.; Moszer, I.; Albertin, A.M.; Alloni, G.; Azevedo, V.; Bertier
 C./Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
 A./Ehrlich, S.D.; Emerson, P.T.; Emtan, K.D.; Erttington, J.; Fabret, C.; Ferrati, E.
 Nature 390, 249-256, 1997
 A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gallier
 A./Iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Koningsstein, G.; Krogh, V.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
 Rieger, M.; Rivola, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, J.; Sato, T.M.; Portetella
 A/Authors: Schleich, S.; Schwoeter, R.; Scoffone, F.; Sekiguchi, Y.; Sekovska, A.; Serot
 A./Reuch, M.; Tanakoshi, A.; Tanaka, T.; Terpestra, P.; Togmon, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
 A/Authors: Yoshikawa, H.F.; Zumeitein, E.; Yoshikawa, H.; Danchin, A.
 A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A/Reference number: A69580; MUID:98044033; PMID:9384377
 A/Accession: B69652
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-132, 'L', 134-212 <KUN>

A/Cross-references: GB:299105; GB:AL009126; NID:9262457; PIDN:CB12064.1; PID:el182222;
 A/Experimental source: strain 168
 C/Genetics:
 A/Gene: *lipA*

Query Match 30.3%; Score 50; DB 2; Length 212;
 Best Local Similarity 38.9%; Pred. No. 8.9;
 Matches 14; Conservative 7; Mismatches 13; Indels 2; Gaps 2;

QY 1 MKFNKKRVAATFALFVSFFITSSQDAQAERS 35
 Db 1 MKFVKRIALVTILMLSVTLFAL-OPSAKAAEHN 35

RESULT 14
 C82242
 oligopeptide ABC transporter, permease protein VC1092 [imported] - *Vibrio cholerae* (strain
 C/Species: *Vibrio cholerae*
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C/Accession: C82242
 R/Heidelberg, J.F.; Bisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwim, M.L.; Dodson, R.J.;
 Chardson, D.; Brimolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoti, I.; Sellers, P.
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A/Reference number: A82035; MUID:20406833; PMID:10952301
 A/Accession: C82242
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-306 <HEI>
 A/Cross-references: UNIPROT:Q9K113; GB:AE004190; GB:AE003852; NID:9655559; PIDN:AAF94253
 A/Experimental source: serogroup O1, strain N16961; Biotype El Tor
 C/Genetics:
 A/Gene: VC1092
 A/Map position: 1
 C/Superfamily: oligopeptide permease protein oppB

Query Match 30.3%; Score 50; DB 2; Length 306;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 12; Conservative 4; Mismatches 6; Indels 2; Gaps 1;

QY 1 MKFNKKRV--AATFALFVSFF 22
 Db 2 LKFIARKIFEAIPMTVLITISFF 25

RESULT 15
 D70141
 oligopeptide ABC transporter, permease protein (oppC-1) homolog - *Lyme disease spirochete*
 C/Species: *Borrelia burgdorferi* (Lyme disease spirochete)
 C/Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
 C/Accession: D70141
 R/Frazer, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White,
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
 Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997
 A/Authors: Smith, H.O.; Venter, J.C.

A>Title: Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*.
 A/Reference number: A70100; MUID:98065943; PMID:9403685
 A/Accession: D70141
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-349 <LKE>
 A/Cross-references: UNIPROT:O51310; GB:AE001140; GB:AE000783; NID:92688233; PIDN:AA6672;
 A/Experimental source: strain B31
 C/Superfamily: oligopeptide permease protein oppB

Query Match 30.3%; Score 50; DB 2; Length 349;
 Best Local Similarity 37.5%; Pred. No. 14;
 Matches 15; Conservative 5; Mismatches 8; Indels 12; Gaps 1;

QY 6 KVAATATFALFVSFFIT-----SSQDAQAER 33
 Db 207 ERSIIIGLFALFVSWLTVARVGRQVSLSSSEPTQAK 246

Search completed: November 7, 2004, 17:57:02
 Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 7, 2004, 14:58:19 ; Search time 101 Seconds
(without alignments)
199,387 Million cell updates/sec

Title: US-09-982-531-1

Perfect score: 165

Sequence: 1 MKFNKKRVAIATFIALIFVSFPTISSQDAQAARS 35

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing filter 45 summaries

Database :

1: uniprot_prot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	130.5	79.1	515	2	Q9RLV2
2	60	36.4	183	2	Q8DRT5
3	59.5	36.1	224	2	Q8V702
4	57	34.5	308	2	Q6YR28
5	57	34.5	308	2	BAD04275
6	57	34.5	578	2	O6LQJ9
7	57	34.5	578	2	CAG20467
8	56.5	34.2	411	1	XYNA_BUTFI
9	56.5	34.2	700	2	Q8ES52
10	54.5	33.0	224	2	Q8V829
11	54	32.7	215	2	Q8VU78
12	54	32.7	275	2	Q891M8
13	54	32.7	370	2	Q93NP8
14	53.5	32.4	447	2	Q9KXJ3
15	53	32.1	252	2	Q81PY2
16	53	32.1	252	2	AAT31773
17	53	32.1	330	2	Q9BKJ9
18	53	32.1	330	2	Q72FW1
19	53	32.1	330	2	AAS94586
20	53	32.1	503	2	Q739A0
21	53	32.1	503	2	Q81E28
22	53	32.1	503	2	AAS41162
23	53	32.1	550	1	USHA_SALPU
24	53	32.1	550	1	USHA_SALTY
25	53	32.1	550	2	Q9R2Z2
26	53	32.1	550	2	Q9RN38
27	53	32.1	550	2	Q9RN39
28	53	32.1	550	2	Q9RN40
29	53	32.1	550	2	Q9RN41
30	52.5	31.8	257	2	Q7CCJ3
31	52.5	31.8	388	2	Q73PV0

32	52.5	31.8	388	2	AAS1189	Aas1189 treponema
33	52.5	31.8	521	2	Q92DS3	Q92ds3 listeria in
34	52.5	31.8	736	2	Q9W1Y0	Q9w1y0 drosophila
35	52.5	31.8	736	2	AAM71132	Aam71132 drosophila
36	52	31.5	141	2	Q972J0	Q972j0 bacillofus
37	52	31.5	208	2	Q6HAN2	Q6han2 bacillus th
38	52	31.5	436	2	Q8KEF1	Q8kef1 chlorobium
39	52	31.5	622	2	O6LS49	O6ls49 photobacter
40	52	31.5	622	2	CAG19877	Cag19877 photobact
41	52	31.5	919	2	Q8TPR8	Q8tp8 methanobact
42	52	31.5	928	2	Q6GB45	Q6gb45 staphylococ
43	52	31.5	933	2	Q53653	Q53653 staphylococ
44	52	31.5	935	2	Q932C5	Q932c5 staphylococ
45	52	31.5	946	2	Q8NXJ1	Q8nxj1 staphylococ

ALIGNMENTS

```

RESULT 1
Q9RLV2 PRELIMINARY; PRT; 515 AA.
AC Q9RLV2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Secreted protein precursor (fragment).
GN Name=310;
OS Lactococcus lactis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1358;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MG1614;
RX MEDLINE=20184754; PubMed=10721729;
RA Ravn P., Arnau J., Maden S.M., Vrang A., Israelien H.;
RT "The development of Tnuc and its use for the isolation of novel
RT secretion signals in Lactococcus lactis.";
RL Gene 242:347-356 (2000).
DR EMBL; AJ238086; CAB61244.1; -.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR008970; Cna_B.
DR InterPro; IPR008985; Cna_Like_g1.
DR Pfam; PF05738; Cna_B; 1.
KW Signal.
FT CHAIN 1 34 Potential.
FT NON_TER 35 515 secreted protein.
SQ SEQUENCE 515 AA; 56591 MW; 2A9C67842CDCA3AB CRC64;

Query Match 79.1%; Score 130.5; DB 2; Length 515;
Best Local Similarity 86.8%; Pred. No. 1,1e-09;
Matches 33; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 MKFNKKRVAIATFIALIFVSFPTISS-QDAQ--AARS 35
DB 1 MKFNKKRVAIATFIALIFVSFPTISSIQDNQTAELS 38

RESULT 2
Q8DRT5 PRELIMINARY; PRT; 183 AA.
AC Q8DRT5;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein SMU_1442c.
OS OrderedCusNames=SMU_1442c;
OC Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MERLIN=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
RA Ajdic D.U., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
RA Ferretti J.U.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL; AF014977; AAN59102.1; -.
DR InterPro; IPR009736; DUF1307.
DR Pfam; PF06998; DUF1307.1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 183 AA; 20583 MW; 41D59B76F448D9 CRC64;

Query Match
Best Local Similarity 36.4%; Score 60; DB 2; Length 183;
Matches 11; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 MKFNKRVATATFIALIFVSFFTTISQDAQAERS 35
Db 5 IRFVKKTLSTVLGLITVAFTLTKVTSANSS 39

RESULT 3
Q8VT02 PRELIMINARY; PRT; 224 AA.
AC Q8VT02;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Agc (Fragment).
GN Name=agc;
OS Staphylococcus capitis subsp. capitis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
ON NCBI_TaxID=72758;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCM2734;
RA Dufour P., Jarraud S., Vandenesch F., Benito Y., Bes M., Greenland T.,
RA Etienne J., Lina G.;
RT "Genetic variability of the agr locus in staphylococcus species.";
RL J. Bacteriol. 0:0-0(2002).
DR EMBL; AF346715; AAL65813.1; -.
FT NON_TER 224
SQ SEQUENCE 224 AA; 26348 MW; D790A3EDC5A638B4 CRC64;

Query Match
Best Local Similarity 36.1%; Score 59.5; DB 2; Length 224;
Matches 13; Conservative 7; Mismatches 13; Indels 1; Gaps 1;

QY 1 MKFNKRVATATFIALIFVSFFTTISQDAQAER 33
Db 141 LSKKKYITITIVLLSFVFFYIISQDMSASD 174

RESULT 4
Q6YR28 PRELIMINARY; PRT; 308 AA.
AC Q6YR28;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE ABC-type dipeptide/oligopeptide transport system, permease
DE component.
GN Name=dppB; OrderedLocuNames=PAM190;
OS Onion yellows phytoplasma.
OC Bacteria; Firmicutes; Mollicutes; Acholeplasmatales;
OC Acholeplasmataceae; Candidatus Phytoplasma.
ON NCBI_TaxID=100379;
RX [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=OY-M;
RX PubMed=14661021;
RA Oshima K., Kakizawa S., Nishigawa H., Jung H.-Y., Wei W., Suzuki S.,
RA Araashida R., Nakata D., Miyata S.-I., Ugaki M., Namba S.;
RT "Reductive evolution suggested from the complete genome sequence of a
RT plant-pathogenic phytoplasma.";
RL Nat. Genet. 36:27-29(2004).
CC -1- FUNCTION: Part of a binding-protein-dependent transport system.
CC Probably responsible for the translocation of the substrate across
CC the membrane (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the binding-protein-dependent transport
CC system permease family.
DR EMBL; AP006628; BAD04275.1; -.
DR InterPro; IPR000515; BPD transp.
DR Pfam; PF00528; BPD transp.1; 1.
DR PROSITE; PS50928; ABC_TM1.1.
KW Complete proteome; Transmembrane; Transport.
SQ SEQUENCE 308 AA; 34832 MW; C82CA02F3261C0C7 CRC64;

Query Match
Best Local Similarity 34.5%; Score 57; DB 2; Length 308;
Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 FNKKRVATATFIALIFVSFFTTI 24
Db 6 FKKSLYTITLFTIVIFISFPTM 27

RESULT 5
BAD04275 PRELIMINARY; PRT; 308 AA.
ID BAD04275
AC BAD04275;
DT 02-MAR-2004 (TREMBLrel. 27, Created)
DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBLrel. 27, Last annotation update)
DE ABC-type dipeptide/oligopeptide transport system, permease
DE component.
GN DPPB OR PAM190.
OS Onion yellows phytoplasma.
OC Bacteria; Firmicutes; Mollicutes; Acholeplasmatales;
OC Acholeplasmataceae; Phytoplasma.
ON NCBI_TaxID=100379;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OY-M;
RX PubMed=14661021;
RA Oshima K., Kakizawa S., Nishigawa H., Jung H.-Y., Wei W., Suzuki S.,
RA Araashida R., Nakata D., Miyata S.-I., Ugaki M., Namba S.;
RT "Reductive evolution suggested from the complete genome sequence of a
RT plant-pathogenic phytoplasma.";
RL Nat. Genet. 36:27-29(2004).
DR EMBL; AP006628; BAD04275.1; -.
SQ SEQUENCE 308 AA; 34832 MW; C82CA02F3261C0C7 CRC64;

Query Match
Best Local Similarity 34.5%; Score 57; DB 2; Length 308;
Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 FNKKRVATATFIALIFVSFFTTI 24
Db 6 FKKSLYTITLFTIVIFISFPTM 27

RESULT 6
Q6LQJ9 PRELIMINARY; PRT; 578 AA.
ID Q6LQJ9
AC Q6LQJ9;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Hypothetical amidase.
GN OrderedLocuNames=PBPRA2004;

```

```

OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=74109;
RN [1]
RP SEQUENCE FROM N.A.
RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,
RA Castaro A., Malacrida G., Simonati B., Cannata N., Bartlett D.,
RA Valle G.;
RT "Genome analysis of Photobacterium profundum reveals the complexity of
RT high pressure adaptations."
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR378669; CAG20407.1; -.
DR InterPro; IPR009070; PGBD_11ke.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 578 AA; 66294 MW; C1CDB7F1DC67AB50 CRC64;

Query Match 34.5%; Score 57; DB 2; Length 578;
Best local Similarity 50.0%; Pred. No. 22;
Matches 12; Conservative 3; Mismatches 9; Indels 0; Gaps 0

QY 3 FNKKRVAIATFIALIFVSPFTISS 26
Db 2 FNHKRIVIGRFSTLVALSFFISS 25

RESULT 7
CAG20407 PRELIMINARY; PRT; 578 AA.
AC CAG20407;
DT 10-MAY-2004 (TrEMBLrel. 27, Created)
DT 10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT 10-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical amidaase.
GN PPRA2004.
OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=74109;
RN [1]
RP SEQUENCE FROM N.A.
RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,
RA Castaro A., Malacrida G., Simonati B., Cannata N., Bartlett D.,
RA Valle G.;
RT "Genome Analysis of Photobacterium profundum reveals the complexity of
RT high pressure adaptations."
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR378669; CAG20407.1; -.
KW Hypothetical protein.
SQ SEQUENCE 578 AA; 66294 MW; C1CDB7F1DC67AB50 CRC64;

Query Match 34.5%; Score 57; DB 2; Length 578;
Best local Similarity 50.0%; Pred. No. 22;
Matches 12; Conservative 3; Mismatches 9; Indels 0; Gaps 0

QY 3 FNKKRVAIATFIALIFVSPFTISS 26
Db 2 FNHKRIVIGRFSTLVALSFFISS 25

RESULT 8
XVNA BUTFI STANDARD; PRT; 411 AA.
AC P23551;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)

```

DE	Endo-1, -beta-xylanase A precursor (EC 3.2.1.8) (Xylanase A) (1,4-
DE	beta-D-xylan xylanhydrolase A).
GN	Name=xynA;
OS	Butyrivibrio fibrisolvens.
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae;
OC	Butyrivibrio.
OX	NCBI_TaxID=831;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=49;
RX	MEDLINE=90330526; PubMed=2198249;
RA	Mannarelli B.M., Evans S., Lee D.;
RT	"Cloning, sequencing, and expression of a xylanase gene from the
RT	aerobic ruminal bacterium Butyrivibrio fibrisolvens.";
RL	J. Bacteriol. 172:4247-4254(1990).
CC	-1- FUNCTION: B.fibrisolvens is located in the rumen of ruminant
CC	animals, where it contributes to the animal's digestion of plant
CC	material by hydrolyzing hemicellulose with its xylanases.
CC	-1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xyloridic
CC	linkages in xyans.
CC	-1- PATHWAY: Xylan degradation.
CC	-1- SIMILARITY: Belongs to cellulase family F (family 10 of glycosyl
CC	hydrolases).
DR	PIR; A37755; A37755.
DR	HSP; P56588; 1831.
DR	InterPro; IPR001000; Glyco_hydr_10.
DR	InterPro; IPR01031; Multisaaem_cyt.
DR	Pfam; PF00331; Glyco_hydr_10_1.
DR	PRINTS; PR00134; GLHYDRLASE10.
DR	SMART; SM00633; Glyco_10; 1.
DR	PROSITE; PS00591; GLYCOSYL_HYDROL_F10; 1.
KW	Glycosidase; Hydrolase; Signal; Xylan degradation.
FT	SIGNAL 1 33 Potential.
FT	CHAIN 34 411 Endo-1,4-beta-xylanase A.
FT	ACT_SITE 201 201 Proton donor (By similarity).
FT	ACT_SITE 311 311 Nucleophile (By similarity).
SQ	SEQUENCE 411 AA; 46659 MW; 4EAAc50D8979DD5 CRC64;
Query Match	34.2%; Score 56.5; DB 1; Length 411;
Best Local Similarity	43.8%; Pred. No. 19;
Matches 11; Conservative 6; Mismatches 11; Indels 1; Gaps 11	
OY	2 KFN-KKRVATPTALIFVSFFTISODAOQA 32 : : : : : 3 KFKIRKLMAVLLALVFSTFFMWSKVDAANA 34
Db	
RESULT 9	
ID	Q8ES52 PRELIMINARY; PRT; 700 AA.
Q8ES52	
AC	Q8ES52;
DT	01-MAR-2003 (TREMBLrel. 23, Created)
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE	Chitinase (EC 3.2.1.14).
GN	OrderedLocustNames=OB0791.
OS	Oceanobacillus iheyensis.
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX	NCBI_TaxID=182710;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=HTE831;
RX	MEDLINE=2220767; PubMed=12325376;
RA	Takami H., Takaki Y., Uchiyama T.;
RT	"genome sequence of Oceanobacillus iheyensis isolated from the Iheya
RT	Ridge and its unexpected adaptive capabilities to extreme
RT	environments.";
RL	Nucleic Acids Res. 30:3927-3935(2002).
DR	HMP; AP004595; BACI2747.1; --
DR	HSSP; Q13211; 1LG2
DR	GO; GO:0009986; Cell surface; IEA.
DR	GO; GO:0004568; F.chitinase activity; IEA.
DR	GO; GO:0016798; F.hydrolase activity; acting on glycosyl bonds; IEA.

```

DR GO:0005975; P:carbohydrate metabolism; IEA.
DR GO:0008152; P:metabolism; IEA.
DR InterPro: IPR011583; Chitinase_11.
DR InterPro: IPR001223; Glyco_hydro_18.
DR InterPro: IPR001579; Glyco_hydro_18A5.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00704; Glyco_hydro_18_1.
DR Pfam: PF00746; Gram_pos_anchor_1.
DR ProDom: PD000471; Chitinase_11; 1.
DR SMART: SM00636; Glyco_18; 1.
DR PROSITE: PS01095; CHITINASE.18; 1.
DR PROSITE: PS50847; GRAM_POS_ANCHORING.1.
KW Complete proteome; Glycosidase; Hydrolase.
SQ SEQUENCE 700 AA; 77273 MW; 56E21B4505E75B7F CRC64;

Query Match
Best Local Similarity 33.3%; Score 56.5; DB 2; Length 700;
Matches 12; Conservative 11; Mismatches 10; Indels 3; Gaps 1;

OY 1 MKFNKRVATFT---ALIFVSFTTSSQDAQA 33
DB 3 LKRRKSLVSSFLMLFCLVFMSPFVANSKVQAQ 38

RESULT 10
OQVS29 PRELIMINARY; PRT; 224 AA.
AC OQVS29;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE AgtC (Fragment).
GN Name=agtC;
OC Staphylococcus capitis subsp. capitis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
CX NCBI_TaxID=72758;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SN910309;
RA Duteour P., Jarraud S., Vandenesch F., Benito Y., Bes M., Greenland T.,
RA Etienne J., Lina G.;
RT "Genetic variability of the agt locus in staphylococcus species.";
RL J. Bacteriol. 0:0-(2002).
DR EMBL: AF346716; AAL55816.1; -.
DR NCBI_TaxID=224
SQ SEQUENCE 224 AA; 26107 MW; 11E3624D944F14C CRC64;

Query Match
Best Local Similarity 33.0%; Score 54.5; DB 2; Length 224;
Matches 13; Conservative 4; Mismatches 14; Indels 1; Gaps 1;

OY 1 MKFNKRVATFTFTALIFVSFTTSSQDAQA 31
DB 141 LSLNKKYITITITIVLFSFVFYIISQDLSA 172

RESULT 11
OQVU78 PRELIMINARY; PRT; 215 AA.
AC OQVU78;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Lipase.
OS Bacillus sp. B26.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CX NCBI_TaxID=117494;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B26;
RA Kim M.H., Kim H.K., Oh T.K., Lee J.K.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF232707; AAL36938.1; -.

```

```

DR HSP: P37957; 116W.
DR GO:0003824; F:catalytic activity; IEA.
DR GO:0016298; F:lipase activity; IEA.
DR GO:0016042; P:lipid catabolism; IEA.
DR InterPro: IPR002918; Lipase_2.
DR InterPro: IPR000379; Ser esters.
DR Pfam: PF01674; Lipase_2_1.
SQ SEQUENCE 215 AA; 22959 MW; 460397D7026A789B CRC64;

Query Match
Best Local Similarity 32.7%; Score 54; DB 2; Length 215;
Matches 12; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

OY 1 MKFNKRVATFTALIFVSFTTSSQDAQAERS 35
DB 4 IRFKKSLQIVLALVLSIAFLQPKKAAAEHN 38

RESULT 12
OQ91M8 PRELIMINARY; PRT; 275 AA.
AC OQ91M8;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Putative glycine-betaine binding permease protein.
GN OrderedLocustNames=CTC02341;
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
CX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Maasachuetts / E88;
RX MEDLINE=22457253; PubMed=12552129;
RA Brueggemann H., Baumeier S., Fricke W.F., Wieszang A.,
RA Decker I., Herzberg C., Martinez-Arias R., Mezkl R., Henne A.,
RA Gottschalk G.;
RT "The genome sequence of Clostridium tetani, the causative agent of
RT tetanus disease.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321 (2003).
CC -1- FUNCTION: Part of a binding-protein-dependent transport system.
CC Probably responsible for the translocation of the substrate across
CC the membrane (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the binding-protein-dependent transport
CC system permease family.
DR EMBL: AB015944; AAC06817.1; -.
DR GO:0016020; C:membrane; IEA.
DR GO:0005215; F:transporter activity; IEA.
DR GO:0006810; P:transport; IEA.
DR InterPro: IPR000515; BPD_transp.
DR Pfam: PF00528; BPD_transp_1; 1.
DR PROSITE: PS50928; ABC_TM1.1.
KW Complete proteome; Transmembrane; Transport.
SQ SEQUENCE 275 AA; 30267 MW; 098FAF2BA45BBD9 CRC64;

Query Match
Best Local Similarity 32.7%; Score 54; DB 2; Length 275;
Matches 12; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

OY 2 KFNKRVATFTALIFVSFTTSSQDAQ 30
DB 62 KLADKRVATFTVGLFPIDSMQLMNQTMQ 90

RESULT 13
OQ3NP8 PRELIMINARY; PRT; 370 AA.
AC OQ3NP8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

```

```

DE Wzy.
GN Name=wzy;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21297197; PubMed=11404020;
RA Wang L., Briggs C.E., Rothmund D., Fratamico P., Luchansky J.B.,
RA Reeves P.R.;
RT "Sequence of the E. coli O104 antigen gene cluster and identification
RT of O104 specific genes.";
RL Gene 270:231-236(2001)
DR EMBL; AF361371; AAK64372.1;
SQ SEQUENCE 370 AA; 43430 MW; E7B867B31E3D68A CRC64;

Query Match 32.7%; Score 54; DB 2; Length 370;
Best Local Similarity 42.3%; Pred. No. 39;
Matches 11; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MKFNKRAVAITFIALIFVSFTTSS 26
Db 154 IKFIEKEPLRYFIATIIASFHLA 179

RESULT 14
Q9KMX3 PRELIMINARY; PRT; 447 AA.
AC Q9KMX3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Na+/H+ antiporter, putative.
GN OrderedLocuNames=VCA0193;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxId=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406633; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gilm M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AB004359; AAF96106.1;
DR PIR; B82490; B82490.
TIGR; VCA0193;
GO; GO:0018021; C:integral to membrane; IEA.
GO; GO:0015385; P:regulation of hydrogen antiporter activity; IEA.
GO; GO:0006885; P:regulation of pH; IEA.
GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR004770; Antiport_nhac.
DR Pfam; PF03553; Na_H_antiporter; 1.
KW Complete proteome.
SQ SEQUENCE 447 AA; 47022 MW; 16988F46BD26D6DC CRC64;

Query Match 32.4%; Score 53.5; DB 2; Length 447;
Best Local Similarity 44.1%; Pred. No. 54;
Matches 15; Conservative 6; Mismatches 12; Indels 1; Gaps 1;

Qy 2 KFNKRAVAITFIALIFVSFTTSSQDAQAERS 35
Db 200 KP-KENIRIALPALALAVVTFSSQTASQPS 232

```

```

RESULT 15
Q81PY2 PRELIMINARY; PRT; 252 AA.
ID Q81PY2; Q6HY52; Q6KS66;
AC Q81PY2; Q6HY52; Q6KS66;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocuNames=BA2658, BAS2476; ORFNames=GBAA2658;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate Porton;
RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
RA Read T.D., Peterson S.N., Tourasse N.J., Baillye L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Rietone J., Wu M.,
RA Kolonay J.F., Beaman M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouiri H.M., Radune D.,
RA Benton J.L., Mahmoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R.T., Redmond C., Thwaitte J.E., White O.,
RA Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate 0581;
RX Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
RA Fraser C.M.;
RT "Bacillus anthracis comparative genomics.";
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Sterne;
RA Brettn T.S., Bruce D., Chailacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Rice H.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017032; AAP26505.1;
DR EMBL; AB017334; AAT31773.1;
DR EMBL; AB017225; AAT54787.1;
TIGR; BA2658;
DR InterPro; IPR010380; DUF975.
DR Pfam; PF06161; DUF975; 1.
KW Hypothetical protein.
SQ SEQUENCE 252 AA; 28338 MW; 9ADE77D331CA9D6 CRC64;

Query Match 32.1%; Score 53; DB 2; Length 252;
Best Local Similarity 61.9%; Pred. No. 39;
Matches 13; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 9 AIAITFIALIFVLFITIGIPDA 29
Db 37 AIAITFIALIFVLFITIGIPDA 57

```

Search completed: November 7, 2004, 17:56:32
Job time : 104 secs

